

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
24 January 2002 (24.01.2002)

PCT

(10) International Publication Number  
**WO 02/06513 A2**

- (51) International Patent Classification<sup>7</sup>: C12Q 1/00 (74) Agent: YANG, Lucy, X.; Intellectual Property Legal Services, Pharmacia & Upjohn Company, 301 Henrietta Street, Kalamazoo, MI 49001 (US).
- (21) International Application Number: PCT/US01/16525
- (22) International Filing Date: 13 July 2001 (13.07.2001) (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/218,118 13 July 2000 (13.07.2000) US  
60/283,880 13 April 2001 (13.04.2001) US
- (71) Applicant (*for all designated States except US*): PHARMACIA & UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): HOMA, Fred, L. [US/US]; 3430 Pine Grove Lane, Kalamazoo, MI 49008 (US). WATHEN, Michael, W. [US/US]; 6474 Pepperidge, Portage, MI 49002 (US). HOPKINS, Todd, A. [US/US]; 744 Sarah Street, Galesburg, MI 49053 (US). THOMSEN, Darrel, R. [US/US]; 6916 Willson Drive, Kalamazoo, MI 49009 (US).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/06513 A2

(54) Title: A METHOD FOR TREATING HERPES VIRUSES

(57) Abstract: The present invention relates to a method for selecting an anti-herpes viral compound and a method for selectively inhibiting herpesvirus in a human host in need of such treatment. The present invention relates to a method for selecting an anti-herpes viral compound and a method for selectively inhibiting herpesvirus in a human host in need of such treatment.

## A METHOD FOR TREATING HERPES VIRUSES

### FIELD OF THE INVENTION

The present invention relates to a method for selecting an anti-herpes viral  
5 compound and a method for selectively inhibiting herpes viruses in a human host in need of such treatment.

### BACKGROUND OF THE INVENTION

The herpesviruses comprise a large family of double stranded DNA viruses. Eight  
10 of the herpes viruses, herpes simplex virus types 1 and 2 (HSV-1 and HSV-2), varicella zoster virus (VZV), human cytomegalovirus (HCMV), Epstein-Barr virus (EBV), and human herpes viruses 6, 7, and 8 (HHV-6, HHV-7, and HHV-8), have been shown to infect humans. Several of these viruses are important human pathogens.

HSV-1 is estimated to affect 100 million people in the U.S. Primary infection of  
15 HSV-1 usually occurs between the ages of one and four. Cold sores, the visible symptom, typically appear at a later age, with 20-45% of the population over the age of fifteen affected (Whitley, Clin. Infect. Dis., 26:541-555, 1998).

Genital herpes (HSV-2) is the second most common sexually transmitted disease, with approximately 22% of the U.S population infected with this virus (Fleming 1997).

20 VZV is the causative agent of chicken pox upon primary infection and can recur in adults as zoster.

EBV results in approximately two million cases of infectious mononucleosis in the U.S. each year. It can also cause lymphomas in immunocompromised patients and has been associated with Burkitt's lymphoma, nasopharyngeal carcinoma, and Hodgkins disease.

25 Infection with HCMV often occurs during childhood and is typically asymptomatic except in immunocompromised patients where it causes significant morbidity and mortality.

HHV-6 is the causative agent of roseola and may be associated with multiple sclerosis and chronic fatigue syndrome. HHV-7 disease association is unclear, but it may  
30 be involved in some cases of roseola. HHV-8 has been associated with Kaposi's sarcoma, body cavity based lymphomas, and multiple myeloma.

These viruses are capable of residing in a latent state within the host. Reactivation of latent virus results from response to environmental stimuli (ex. UV exposure, stress,

etc.). Infections or recurrence can be life threatening in immunocompromised patients such as AIDS or transplant patients where HCMV can result in retinitis, pneumonia, and gastrointestinal disease.

The increased immunocompromised population has created an unmet medical need for antivirals against herpesviruses because current therapies do not have a sufficiently broad spectrum against this family of viruses and/or they have limited utility due to toxicity. The present invention provides a method for selectively inhibiting herpesviruses DNA polymerase with compounds that have broad spectrum activity. The method offers a distinct advantage in the treatment of patients in need, particularly immunocompromised patients at risk of infection or reactivation by many members of the herpesvirus family.

### SUMMARY OF THE INVENTION

The present invention provides a method of selecting compounds that inhibit herpes viruses comprising:

- a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type herpes virus,
- b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant herpes virus which is the same strain of the wild type herpes virus,
- c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
- d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.

In above method, the order of step a and step b are interchangeable.

The present invention further provides a method of selecting compounds that inhibit herpes viruses comprising:

- a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HSV-1,
- b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HSV-1 which is the same strain of the wild type herpes virus,
- c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
- d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.

In above method, the order of step a and step b are interchangeable.

The present invention further provides a method of selecting compounds that inhibit herpes viruses comprising:

- a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HSV-2,

- b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HSV-2 which is the same strain of the wild type herpes virus,
- c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
- d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.

In above method, the order of step a and step b are interchangeable.

The present invention further provides a method of selecting compounds that inhibit herpes viruses comprising:

- a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HCMV,
- b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HCMV which is the same strain of the wild type herpes virus,
- c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
- d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.

In above method, the order of step a and step b are interchangeable.

The present invention further provides a method for selectively treating diseases caused by herpes viruses in a human host comprising administering a compound to a human in need of such treatment wherein said compound inhibits herpes viruses by interaction with the binding domain in the viral DNA polymerase.

The present invention further provides method for selectively inhibiting herpes viruses in a human host comprising administering a compound to a human in need of such treatment wherein  $IC_{50}$  of the compound that inhibits a binding domain mutant herpes virus is at least 3 times greater than  $IC_{50}$  of the compound that inhibits a wild type herpes virus which is the same strain as the mutant herpes virus.

The present invention further provides a compound for treating herpesviral infections in a human host wherein  $IC_{50}$  of the compound that inhibits a binding domain mutant herpes virus is at least 5 times greater than  $IC_{50}$  of the compound that inhibits a wild type herpes virus which is the same strain as the mutant herpes virus.

The present invention further provides a compound for treating herpesviral infections in a human host wherein said compound inhibits the herpesvirus by interacting with the binding domain in the viral DNA polymerase.

The present invention further provides a compound for the inhibiting of herpesvirus DNA polymerases wherein serial passage of a wild type herpes virus in the presence of said

compound results in a change of the wild type HSV-1 polymerase at amino acid 823 from valine to alanine.

The present invention further provides a compound for inhibiting herpesvirus DNA polymerases wherein serial passage of a wild type herpes virus in the presence of said  
5 compound results a change of the wild type HCMV polymerase at amino acid 823 from valine to alanine and at amino acid 824 from valine to leucine.

The present invention further provides a mutant herpesvirus DNA molecule having a nucleotide sequence selected from a group consisting of SEQ.ID.NO. 1; SEQ.ID.NO. 3; SEQ.ID.NO. 5; SEQ.ID.NO. 7; SEQ.ID.NO. 9; and SEQ.ID.NO. 11.

10 The present invention further provides a mutant herpesvirus polymerase amino acid molecule having an amino acid sequence selected from a group consisting of SEQ.ID.NO. 2; SEQ.ID.NO. 4; SEQ.ID.NO. 6; SEQ.ID.NO. 8; SEQ.ID.NO. 10 and SEQ.ID.NO. 12.

#### BRIEF DESCRIPTION OF THE DRAWINGS

15 Figure 1 – examples of 4-oxo-DHQ and 4-oxo-DHTP compounds.

Figure 2 – Herpesvirus' polymerases amino acid conserved region.

Figure 3 – Recovered virus after serial passage of HSV-1 in presence of 20  $\mu$ M of compound No. 17.

Figure 4 – Comparison of Wild HSV-1 and HSV-2 herpesvirus DNA polymerase  
20 amino acid sequences aligned by amino acid homology. (Seq. No: 14-19)

Figure 5 – Mutant Herpes Virus DNA and amino acid sequence list. (Seq. No: 1-12)

Figure 6 – Wild HCMV herpesvirus DNA polymerases amino acid sequence. (Seq. No 13)

#### 25 DETAILED DESCRIPTION OF THE INVENTION

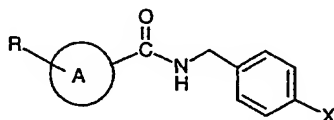
A key enzyme in the replication of all herpesviruses is the virus-coded DNA polymerase. Most of the currently available anti-herpes drugs target the viral DNA polymerase. Drugs such as Foscarnet acts by direct inhibition of the viral polymerase. These drugs are non-nucleoside inhibitors of herpesvirus DNA polymerases. Others such as the  
30 nucleoside analogs, Acyclovir, Penciclovir and Ganciclovir must first be phosphorylated to the monophosphate forms by virus encoded kinases and, further phosphorylated to triphosphate by cellular enzymes before they are active inhibitors. The triphosphate forms of these nucleoside analogs inhibit polymerases by competing with the binding of natural

triphosphates and their subsequent insertion into growing DNA strands. These drugs are known as nucleoside inhibitors of herpesvirus DNA polymerases.

One of the limitations of the currently available drugs is that they are active against only a few of the eight human herpesviruses. For example, Acyclovir and Penciclovir  
5 inhibit HSV and VZV replication but have poor activity against CMV.

In order to identify antiviral compounds that would have the potential to inhibit replication of most of the human herpesviruses, compounds are *in vitro* screened for inhibitors of herpesvirus DNA polymerase activity. Because portions of the amino acid sequence of the polymerases are highly conserved within the herpesvirus family it is  
10 possible to discover small molecules that inhibit herpesvirus polymerases but not cellular DNA polymerases. Using this biochemical approach, several new classes of compounds such as the 4-hydroxyquinoline derivatives (4-HQ), 4-oxo-dihydroquinoline derivatives (4-oxo-DHQ) and 4-oxo-dihydrothienopyridine derivatives (4-oxo-DHTP) were discovered as potent, non-nucleoside herpesvirus DNA polymerase inhibitors. *In vitro* polymerase assays  
15 and/or *in vivo* cell culture assays have demonstrated that these compounds inhibit HSV-1, HSV-2, HCMV, VZV, EBV, and HHV-8 replication.

4-Oxo-DHQ and 4-oxo-DHTP are derivatives of formula I



20 wherein ring A is a saturated or unsaturated fused double or triple heterocyclic ring having 1, 2, 3 or 4 heteroatoms selected from group consisting of oxygen, sulfur, or nitrogen; and wherein R and X are the appropriated substitutents, respectively.

Examples of 4-HQ compounds, 4-oxo-DHQ compounds and 4-oxo-DHTP compounds are illustrated in **Figure 1**.

25 Antiviral activity of these examples are shown in Table 1 below. As shown in Table 1, these compounds inhibit HSV-1 and HSV-2 as well or better than the current commercially available drug Acyclovir.

**Table 1**  
**Antiviral Activity of 4-oxo DHQ/4-oxo DTHP Against HSV-1 and HSV-2**

virus	Compound IC <sub>50</sub> (uM)					ACV
	1	2	3	4	5	
HSV-1 KOS	2.0	3.8	3.2	3.2	3.3	3.6
HSV-1 F	2.5	2.3	2.2	2.1	2.6	1.3
HSV-1 DJL	2.5	2.6	1.8	2.2	2.7	1.8
HSV-1 Patton	ND	5.3	7.7	4.3	10	9.3
HSV-2 MS	2.0	2.5	2.8	2.5	2.5	10
HSV-2 35D	ND	5.4	5.0	3.2	8.1	6.0
HSV-2 186	2.0	2.3	3.2	2.3	4.2	>10

5 It has also been discovered that point mutations within the HSV-1 polymerase gene that confer resistance to Acyclovir and other nucleoside analogs do not result in resistance to the 4-HQ, 4-oxo-DHQs or 4-oxo-DHTPs. Serial passage of wild type HSV-1 in the presence of 4-oxo-DHQ results in the isolation of mutants that are highly resistant (>20 fold increase in the IC<sub>50</sub>) to these compounds while retaining sensitivity to nucleoside inhibitors  
 10 such as Acyclovir.

In order to determine the mechanism of action of 4-HQ, 4-oxo-DHQ and 4-oxo-DHTP compounds against herpes viruses, mutants resistant to these compounds are isolated by serial passage of the virus in the presence of a 4-oxo-DHQ compound. Sequencing analysis of HSV-1 and HSV-2 strains resistant to the 4-oxo-DHQ identifies that HSV-1  
 15 (KOS strain) polymerase protein and its homologous HSV-2 have a conserved region (a binding domain), which is a critical contact point for these compounds. While amino acid numbering of the DNA polymerase may vary between strains of HSV-1 and HSV-2, this binding domain encompassing the HSV-1 (KOS) strain amino acid 823 is highly conserved in herpesviruses and can be identified by aligning the homologous amino acids of this  
 20 domain as shown in Fig 2.

In HSV-1 and HSV-2 strains resistant to the 4-oxo-DHQ and similar compounds, a change of valine to an alanine at the binding domain provides full resistance.

In the HSV-1 DNA polymerase, resistance is also found when a valine changes to methionine at amino acid 823 but only when accompanied by a second amino acid change.

25 Isolation of HCMV resistant to 4-oxo-DHQ's is found to be very difficult. Comparison of the amino acid sequence of the HSV polymerase (Y-G-F-T-G-V-Q-H-G) and HCMV polymerase (Y-G-F-T-G-V-V-N-G) in the region of amino acid 823 (underlined amino acid) shows that there is a second valine at position 824 in the HCMV

polymerase. In vitro assay using mutant HCMV polymerases demonstrates that full resistance to the 4-oxo-DHQs requires changes at both amino acids 823 (a valine to alanine) and 824 (a valine to leucine). A HCMV polymerase gene containing V823A and V824L mutations is used in marker rescue experiments to generate a viral mutant. This mutant has an  $IC_{50}$  approximately 7-fold above that of wild-type HCMV.

The HSV-1, HSV-2 and HCMV mutants are also found to be resistant to other non-nucleoside inhibitors such as the 4-oxo-DHTP and similar compounds. However, when the binding domain mutants (e. g. HSV-1 V823A, HSV-2-MS V826A, HSV-2-186 V828A, and HCMV V823A/V824L mutants) are tested in plaque reduction assays against a series of nucleoside polymerase inhibitors and the non-nucleoside inhibitor such as Foscarnet, replication of the mutants is found to be inhibited by all of the currently marketed anti-herpes polymerase inhibitors tested.

These studies demonstrate that certain non-nucleosides like 4-HQ, 4-oxo-DHQ and 4-oxo-DHTP compounds bind to a different site on the herpes polymerase than the nucleoside inhibitors and Foscarnet. The valine at the binding domain is conserved in the DNA polymerases of six of the eight human herpesviruses and several animal herpesviruses, and appears to play a critical role in the antiviral activity of the 4-HQ, 4-oxo-DHQ and 4-oxo-DHTP compounds. (See Figure 2)

Since mutation at the binding domain negates these non-nucleoside inhibitors' activities, compounds could be tested against wild type polymerases and the mutant polymerases to establish the probability of similar binding. We refer to this property of compounds as interaction with the binding domain. Since compounds that interact with the binding domain have exhibited broad-spectrum activity against herpesviruses, this invention provides a method for selecting compounds to treat individuals such as immunocompromised patients who are afflicted with multiple herpesvirus infections.

### **Definitions**

The term "wild-type" refers to a gene or gene product which has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "normal" or "wild-type" form of the gene.

In contrast, the term "mutant" refers to a gene or gene product which displays modifications in sequence and or functional properties (i.e., altered characteristics) when

compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

IC<sub>50</sub> refers to concentration of a drug that inhibits virus growth by 50%.

5 Wild type HSV-1 and HSV-2 strains are listed in **Figure 4**.

Wild type HCMV is listed in SEQ. ID. NO.13.

The term "Iudr" refers to antiviral drug Iododeoxyuridine.

The term "Bvdu" refers to antiviral drug Bromovinyldeoxyuridine.

The term "ACV" refers to antiviral drug Acyclovir.

10 The term "AraC" refers to antiviral drug Arabinosylcytidine.

The term "AraT" refers to antiviral drug Arabinosylthymine.

The term "AraA" refers to antiviral drug Arabinosyladenine.

The term "GCV" refers to antiviral drug Ganciclovir.

The term "CDV" refers to antiviral drug Cidofovir.

15 The term "PFA" refers to antiviral drug Foscarnet.

The term "binding domain" refers to a conserved region in herpesvirus DNA polymerases. The herpesvirus DNA polymerases have seven (7) conserved regions. The binding domain is within the third conserved region (see Figure 2). When the binding domain contacts with an inhibitor, at least one amino acid in the binding domain mutates and provides the resistance. In general, the binding domain is at an amino acid sequence position 818-829 of the HSV-1 DNA polymerase or the homologous region in other herpes virus DNA polymerases (see Figure 2).

The term "a binding domain mutant herpes virus" refers to a herpes virus containing a binding domain mutation.

25 More specifically, the binding domain in HSV-1 strains, KOS, F, DJL and Patton are at amino acid sequence position 823. The binding domain in HSV-2 MS-M1 strain is at amino acid sequence position 826. The binding domain in HSV-2 186 strain is at amino acid sequence position 828. The binding domain in HCMV AD 169 strains is at amino acid sequence position 823-824.

30 The term "XxxxY" refers to an amino acid sequence position xxx, a single amino acid X in wild type is changed to an amino acid Y.

For example, the term "V823A" refers to an amino acid sequence position 823, a Valine found in wild type is changed to alanine in mutant strain.

The term "V824L" refers to an amino acid sequence position 824, a Valine found in wild type is changed to Leucine in mutant strain.

The term "V826A" refers to an amino acid sequence position 826, a Valine found in wild type is change to alanine in mutant strain.

- 5 The term "V828A" refers to an amino acid sequence position 828, a Valine found in wild type is change to alanine in mutant strain.

A table of amino acids and their representative abbreviations, symbols and codons is set forth below in the following Table.

10

Amino acid	Abbrev.	Symbol	Codon(s)					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

## MATERIALS AND METHODS

### Cell and Viruses

- 15 African green monkey kidney cells (Vero) and human foreskin fibroblast cells (HFF) and herpes viruses can be obtained from the American Type Culture Collection (ATCC). Media is defined as Dulbecco's modified Eagle media (DMEM) containing 10% fetal bovine serum (FBS) and supplemented with antibiotics. Cells are maintained in media at 37°C in a humidified atmosphere of 5% CO<sub>2</sub>. HSV-1 strains F, Patton and DJL, HSV-2 strains MS, 35D and 186, and HCMV strain AD169 are used in these studies. Strain DJL is
- 20 a clinical isolate of HSV-1 isolated in our lab from a primary oral lesion.

**Measuring IC<sub>50</sub> of a Compound of Interest That Inhibits Herpes Viruses**

**Preparation of Virus Stocks:** HSV-1 and HSV-2 stocks are grown in Vero cells. HCMV stocks are grown in HFF cells. Approximately 1 ml of media containing sufficient virus to infect approximately 0.1% to 1% of the cells (multiplicity of infection of 0.001 to 5 0.01 PFU/cell) is added to a T-150 cell culture flask containing a confluent monolayer of cells. The cells are incubated at 37°C for approximately 1 hour. Approximately 50 ml of media is then added to the flask and the cells are incubated at 37°C until viral cytopathic effect (cpe) is apparent in 100% of the cells. The flask is then placed at -80°C for at least 30 min. The flask containing frozen media and cells is placed in a 37°C water bath until the 10 media is thawed. This process disrupts the cells and releases virus into the media. 1 ml aliquots of media containing virus are dispensed into tubes and stored at -80°C. These aliquots of media containing virus are referred to as virus stocks.

**Titrating Virus Stocks:** Aliquots of virus are thawed at 37°C and serially diluted (10 fold dilutions) in media. 0.1 ml of each dilution of virus is placed in a single well of 24- 15 well cell culture dish containing a confluent monolayer of cells (Vero cells for HSV-1 and HSV-2, HFF cells for HCMV) and incubated at 37°C for 1 h. The virus inoculum is then removed and 1 ml of media containing 0.8% carboxymethylcellulose (CMC) is added to each well of the dish. The dish is incubated at 37°C for approximately 2-3 days (HSV-1 and HSV-2) or 6-9 days (HCMV) to allow sufficient growth of virus to form plaques in the 20 cell monolayer. Plaques can be observed and counted microscopically or by staining the cells with 0.1% crystal violet in 20% ethanol. The virus titer which is expressed as plaque forming units (PFU) per ml is obtained by counting the plaques in a well and correcting for the dilution of the viral inoculum.

**Plaque Reduction Assays:** Antiviral activity of compounds against herpesviruses such as 25 HSV-1, HSV-2, or HCMV can be measured using plaque reduction assays. 0.1 ml of media containing approximately 50 PFU of virus is added to each well of a 24-well cell culture dish containing a confluent monolayer of cells (Vero cells for HSV-1 and HSV-2, HFF cells for HCMV). Compounds are dissolved in 100% DMSO and diluted in 100% DMSO as 200x stocks of the desired final drug concentration. Typically 5-6 two-fold dilutions are 30 prepared for each compound. Dilutions of compounds are then added to media containing 0.8% CMC resulting in a final 1x drug concentration. After the virus-infected cells have incubated for 1 h at 37°C, the virus inoculum is removed and 1 ml of media containing 0.8% CMC and the various concentrations of compound is added to each well of the dish.

The dish is incubated at 37°C for approximately 2-3 days (HSV-1 and HSV-2) or 6-9 days (HCMV) to allow sufficient growth of virus to form plaques in the cell monolayer. Plaques can be observed and counted microscopically or by staining the cells with 0.1% crystal violet in 20% ethanol. Virus inhibition is determined for each drug concentration by  
5 comparing the number of plaques in drug-containing wells to control wells that did not contain drug. Antiviral activity of a compound is expressed as the concentration of compound predicted to reduce the number of plaques in a well by 50% (IC<sub>50</sub>). The IC<sub>50</sub> values are calculated by plotting the per cent inhibition vs. concentration of compound using EXCEL software for linear regression.

10

#### **Selection of 4-oxo-DHQ resistant HSV-1 and HSV-2**

Vero cells are plated out at a density of  $3.5 \times 10^5$  cells per well in a six well tissue culture plate. Cells are infected with HSV-1 KOS at a multiplicity of infection (moi) of 0.1 pfu/cell and 1 h post infection the cells are overlaid with 3 ml media containing 20  
15 uM of a 4-oxo-DHQ. Cultures are incubated for 20 h at 37°C, freeze/thawed to release cell-associated virus, and 0.1 ml of culture is used to infect a new monolayer of Vero cells (one passage). Serial passage is repeated seven times in the presence of 20 uM drug. Virus isolates are then plaque purified three times prior to preparation of stocks. Virus recovered from each passage in the presence of compound No. 17 is shown in Figure 3. 4-oxo-DHQ  
20 resistant HSV-1 and HSV-2 may also be selected by the marker transfer method described below using wild-type HSV DNA and the corresponding mutant HSV polymerase gene.

#### **Marker Transfer of a HCMV Mutation**

A plasmid containing the wild-type HCMV polymerase gene is modified to contain  
25 the V823A or V823A and V824L mutations using a site-directed mutagenesis Kit (Stratagene Corp.) and following the manufacturer's protocol. HFF cells are plated into T25 tissue culture flasks to achieve 80% confluency at the time of the transfection. Wild type HCMV AD169 DNA and plasmid DNA containing the mutant HCMV polymerase gene are mixed at a ratio of 1:2 (2ug of viral DNA to 4 ug of plasmid DNA). DNA's are  
30 transfected using superfect transfection reagent according to methods recommended by the manufacturer (Quiagen Inc.). Cells are harvested five days posttransfection, freeze-thawed to release virus and half of the sample is used to infect HFF cell monolayers. Cells are overlaid with media containing 20 uM 4-oxo-DHQ compound 2 (see Figure 1). Serial

passage is repeated seven times in the presence of 20 uM compound 2 and virus isolates are then plaque purified three times prior to preparation of viral stock.

#### Isolation of HSV and HCMV viral DNA

- 5 HSV DNA is purified from the cytoplasm of infected Vero cells. Vero cells (50 % confluent) are infected at an multiplicity of 0.01 PFU/cell. At 3-5 days postinfection infected cells (100% cpe) are harvested by centrifugation at 1000 rpm in a Beckman GS-6R centrifuge. The pelleted cells are resuspended in TE buffer and placed on ice for 15 minutes. NP-40 is then added to a final concentration of 0.2% and incubated on ice for a
- 10 further 15 minutes. The cells are centrifuged at 2000 rpm for 10 minutes in a Beckman GS-6R centrifuge. The supernatant is removed and EDTA is added to a final concentration of 20 mM followed by the addition of SDS to a final concentration of 0.3% and proteinase K to a concentration of 50 ug/ml then incubated at 45C for 2 hours. HCMV DNA is isolated by infecting HFF cells (25% confluency) with HCMV at an multiplicity of 0.1 PFU/cell.
- 15 Cells and media are harvested 5-7 days postinfection (100% cpe) and subjected to low speed centrifugation to remove intact cells and cell debris followed by a high speed spin to pellet virus particles (2500 rpm's in a Beckman SW28 rotor for 1 hour). Following incubation of the HSV and HCMV samples, 1.5 volumes of saturated NaI is added to the digested extract and the refractive index is adjusted to 1.434 –1.435. Ethidium bromide is
- 20 added to a final concentration of 50 ug/ml. The samples are loaded into a VTI 50centrifuge tube and spun for 24 hours at 45,000 rpm. The DNA band is harvested extracted three times with n-butanol, then dialyzed against TE buffer followed by a dialysis against 95% ethanol and a final dialysis against TE buffer.

#### 25 DNA Sequencing

- HSV-1, HSV-2 or HCMV viral DNA's are sequenced directly using an ABI377 fluorescence sequencer (Perkin Elmer Applied Biosystems, Foster City, CA) and the ABI BigDye PRISMTM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq FSTM DNA polymerase (PE Applied Biosystems). Each cycle sequencing
- 30 reaction contained about 1.0 ug of purified viral DNA. Cycle-sequencing is performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 98°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times are controlled by a Perkin-Elmer 9700 thermocycler. Extension products are

purified using Centriflex™ gel filtration cartridges (Edge BioSystems, Gaithersburg, MD). Each reaction product is loaded by pipette onto the column, which is then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B table top centrifuge) at 750 x g for 1.5 min at room temperature. Column-purified samples are dried under vacuum for about 40 min and then dissolved in 4 ul of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples are then heated to 90°C for two min, and held at 4°C until loading. 1.5 ul of each sample is loaded into a single well of the ABI377 sequencer. Sequence chromatogram data files from the ABI377 are analyzed with the computer program Sequencher (Gene Codes, Ann Arbor, MI), for assembly of sequence fragments and correction of ambiguous base calls. Generally sequence reads of 600-700 bp are obtained. Potential sequencing errors are minimized by obtaining sequence information from both DNA strands and by re-sequencing difficult areas using primers at different locations until all sequencing ambiguities are removed.

The entire coding region of the polymerase genes from both the parent strains and the resistant viruses are sequenced. The DNA sequencing is done using viral DNA as the template thus avoiding cloning of the polymerase genes. The amino acid sequence of the DNA polymerases of HSV-1 KOS, F, Patton and DJL and HSV-2 MS and 186 are compared in Figure 4. Amino acids that are identical for the six polymerases are shaded in black while regions where amino acid differences are found are shaded in gray. The amino acid sequence of the four HSV-1 polymerases are essentially identical with only a few minor changes noted between the different HSV-1 strains. The majority of amino acid changes are found when the sequences of the HSV-1 and HSV-2 polymerases are compared.

#### **Isolation and Characterization of HSV-1 and HSV-2 Mutants That Are Resistant To the 4-oxo-DHQ's and 4-oxo-DHTP Compounds**

A panel of viruses consisting of four strains of HSV-1 (KOS, F, DJL, Patton) and three strains of HSV-2 (MS, 35D, 186) are tested in a plaque reduction assay against four different 4-oxo-DHQ compounds (# 1, 2, 4, 5 as shown in Figure 1), and one 4-oxo-DHTP compound (# 3 as shown in Figure 1) and against Acyclovir. The six drugs inhibited replication of the seven virus strains with IC<sub>50</sub> values ranging from 2-10 µM (Table 1). In order to select for 4-oxo-DHQ resistant mutants, HSV-1 strains KOS, F, and DJL along with HSV-2 strains 186 and MS are serially passaged in the presence of 20 uM compound

1. Following the seventh passage, 4-oxo-DHQ resistant virus from each strain are plaque purified three times and high-titer stocks are made. All of the resistant HSV mutants grew to high titers in Vero cells, indicating that the mutations in the resistant isolates did not significantly impair their growth. The mutants selected with 4-oxo-DHQ compound 1 exhibited >10 fold increase in IC<sub>50</sub> when tested in a plaque reduction assay against 4-oxo-DHQ compound 1. Data are shown in Table 2.

Table 2

**4-oxo-DHQ Resistant Virus of HSV-1 and HSV-2**

Virus Mutants	Compound 1 IC <sub>50</sub> (uM)	Amino Acid Change in HSV DNA Polymerase
HSV-1 Kos-M1	>20	- V823A
HSV-1 F-M1	>20	- V823A
HSV-1 DJL-M1	>20	-V823A
HSV-2 MS-M1	>20	- V826A
HSV-2 186-M1	>20	- V828A

- \*HSV-1 and HSV-2 isolates grown in the presence of 4-oxo-DHQ select for resistant virus.

DNA sequence analysis of the 4-oxo-DHQ resistant mutants (HSV-1 KOS-M1, HSV-1 F-M1, HSV-1 DJL-M1, HSV-2 186-M1, HSV-2 MS-M1) demonstrated that all five mutants contained a single point mutation of T to C at the binding domain resulting in a Valine to Alanine amino acid change.

**Isolation and Characterization of A HCMV Mutant That Is Resistant to The 4-oxo-DHQ's and 4-oxo-DHTP Compounds**

- In order to select for a 4-oxo-DHQ HCMV resistant mutant, virus (strain AD169) is serially passaged in the presence of 20 uM a 4-oxo-DHQ. Although we could readily select for HSV mutants using this procedure we failed to isolate an HCMV mutant, even when the virus is passaged at low drug concentrations (<5 uM). Comparison of the amino acid sequence of the HSV polymerase, Y-G-F-T-G-V-Q-H-G, and HCMV polymerase, Y-G-F-T-G-V-V-N-G, in the region of amino acid 823 (underlined amino acid) showed that there is a second valine at position 824 in the HCMV polymerase. In order to determine if both valines need to be changed in order to confer resistance to the 4-oxo-DHQ's, *in vitro* polymerase assays are done using mutant HCMV polymerases containing either V823A or V823A plus V824L (Table 3).

**Table 3**  
**HCMV Mutant Polymerase Exhibits Resistance to 4-oxo-DHQ\***

Polymerase	Compound 1 IC <sub>50</sub> (uM)
HCMV (wild)	4.6
HCMV V823A	17.2
HCMV V823A/V824L	42.9

\*Generation of the valine to alanine at amino acid 823 of HCMV results in a 3.5-fold increase in resistance.

\*Mutation of the amino acid from valine to alanine and amino acid 824 from valine to leucine results in an 9-fold increase in resistance, relative to wild type.

The V823A alone resulted in a 3.5-fold increase in the IC<sub>50</sub> while the polymerase with the double amino acid change had nearly 10-fold increase in the IC<sub>50</sub>. In order to isolate an HCMV resistant mutant marker rescue experiments are done. Plasmids containing the mutant polymerase genes are transfected into HFF cells along with wild type HCMV AD169 DNA. The resulting virus is then serially passaged in the presence of 20 uM compound 1 (see figure 1). A 4-oxo-DHQ resistant virus is isolated from marker rescue studies done with the HCMV polymerase gene containing mutations that result in the V823A, V824L amino acid changes, but not with the gene containing V823A change alone. The mutant selected with compound 1 (HCMV AD169-M1) exhibited ~7-fold increase in IC<sub>50</sub> when tested in a plaque reduction assay compared to Ganciclovir and

20 cidofovir which has a  $\leq$  2-fold change in sensitivity (Table 4).

**Table 4**  
**Plaque reduction assay of 4-oxo-DHQ resistant HCMV\***

Drug	HCMV AD169 IC <sub>50</sub> (μM)	HCMV AD169 – M1 IC <sub>50</sub> (μM)
Compound 1	0.7	4.7
Ganciclovir	0.9	1.0
Cidofovir	0.3	0.6

\*Recombination of wild-type HCMV with a polymerase gene containing the valine to alanine at amino acid 823 and the valine to leucine at amino acid 824 allowed for selection of resistant virus with about 7-fold less sensitivity to compound 1.

\*Sensitivity of resistant HCMV virus to Ganciclovir and Cidofovir verifies that the 4-oxo-DHQ's mechanism for inhibiting the polymerase protein is unique

The entire coding region of the HCMV polymerase genes from both the parent strain and the resistant virus are sequenced. The DNA sequencing is again done using viral DNA as the template thus avoiding cloning of the polymerase genes. Comparison of the DNA sequence of the two polymerase genes demonstrated that the resistant mutant  
5 contained two point mutations that resulted in the predicted V823A, V824L amino acid changes. As with the HSV resistant viruses these results demonstrate the critical role of the region encompassing amino acid 823 for inhibition of polymerase activity by these compounds.

10 **Antiviral Activity of Nucleoside and Non-Nucleoside Polymerase Inhibitors Against 4-oxo-DHQ Resistant Mutants**

In order to determine if the 4-HQ binding domain mutations alter the sensitivity of the HSV-1, HSV-2 and HCMV mutants to both non-nucleoside (4-oxo-DHQ's) and nucleoside inhibitors (e.g Acyclovir and ganciclovir) several of the mutants are tested in  
15 plaque reduction assays against a series of non-nucleoside compounds including Foscarnet (PFA), 4-HQ's 4-oxo-DHQ's and 4-oxo-DHTP's (Table 5). The mutants are also tested against a series of nucleoside inhibitors including acyclovir and ganciclovir (Table 5). The activity of these compounds against the mutants is compared to their activity against the wild type strains that are used to isolate the HSV and HCMV mutants. When tested against  
20 a number of 4-HQ's, 4-oxo-DHQ's and 4-oxo-DHTP's and other related classes of compounds all of the drugs are found to inhibit the wild type virus with IC<sub>50</sub> values ranging from <0.1 uM to 30 uM. When these drugs are tested against the resistant viruses they are found to have IC<sub>50</sub> values 5 to 10 fold higher than the parent virus. There is little if any difference in the IC<sub>50</sub> values of the nucleoside compounds and the non-nucleoside PFA  
25 between the wild type and mutant HSV-1, HSV-2, and HCMV viruses. These results demonstrate that the amino acid change in the binding domain (V823A in the HSV-1 polymerase, V826A in the HSV2-MS polymerase, V828A in the HSV2-186 polymerase, and the V823A/V824L changes in the HCMV polymerase) resulted in resistance to the 4-oxo-DHQ's and 4-oxo-DHTP's, which provides further evidence that these classes of  
30 compounds share an affinity for a region we refer to as the binding domain. In contrast, these amino acid changes did not alter the activity of these viruses to other classes of polymerase inhibitors.

Table 5

Antiviral activity of nucleoside and non-nucleoside polymerase inhibitors  
against HSV-1, HSV-2, and HCMV Isolates selected for 4-oxo-DHQ resistance\*

Drug	Plaque Reduction Assay – IC <sub>50</sub> (μM)					
	HSV-2 MS	HSV-2 MS-M1	HSV-1 KOS	HSV-1 KOS-M1	HCMV AD169	HCMV AD169-M1
6	28.8	>50	24.6	>50	5.1	>16
7	8.8	27.9	6.5	>50	0.3	3.4
8	2.3	>50	5.1	>50	<0.1	1.1
9	0.9	48.7	1.9	>50	<0.1	3.1
10	29.2	>50	15.8	>50	1.1	>16
11	3.0	>50	3.1	>50	0.7	3.9
12	0.4	12.5	1.3	>50	0.2	1.1
13	5.3	>50	5.5	<25	2.7	>16
14	1.6	>50	28.4	>50	0.9	18.4
2	1.3	>50	3.3	>50	0.4	4.0
4	2.1	28.4	4.2	>50	0.6	2.1
3	0.8	>50	4.0	>50	1.5	6.2
15	5.9	>50	>50	>50	0.7	7.7
Iudr	5.0	6.1	1.1	0.8	ND	ND
Bvdu	5.8	5.9	2.1	0.1	ND	ND
ACV	2.4	2.8	3.9	4.4	ND	ND
AraC	0.2	0.1	0.2	0.2	ND	ND
AraT	6.6	3.6	11.6	3.6	ND	ND
AraA	10.6	18.2	26.1	27.2	ND	ND
GCVir	ND	ND	ND	ND	0.8	0.8
CDV	ND	ND	ND	ND	0.4	0.3
PFA	ND	ND	ND	ND	38	<20

5 \*HSV-2 MS, HSV-1 KOS, HCMV AD169: wild type strains

\*HSV-2 MS-M1, HSV-1 KOS-M1, HCMV AD169-M1: mutants selected for 4-oxo-DHQ resistance

\*ND – Not Done.

Antiviral compounds identified by the present invention can conveniently be  
10 administered in a pharmaceutical composition containing the compound in combination  
with a suitable excipient, the composition being useful in combating viral infections.  
Pharmaceutical compositions containing a compound appropriate for antiviral use are  
prepared by methods and contain excipients which are well known in the art. A generally  
recognized compendium of such methods and ingredients is Remington's Pharmaceutical  
15 Sciences by E.W. Martin (Mark Publ. Co., 15th Ed., 1975).

Antiviral compounds identified by the present invention and their compositions can  
be administered parenterally (for example, by intravenous, intraperitoneal or intramuscular

injection), topically, orally, or rectally, depending on whether the preparation is used to treat internal or external viral infections.

For oral therapeutic administration, the active compound may be combined with one or more excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 0.1% of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 60% of the weight of a given unit dosage form. The amount of active compound in such therapeutically useful compositions is such that an effective dosage level will be obtained.

The tablets, troches, pills, capsules, and the like may also contain the following: binders such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, fructose, lactose or aspartame or a flavoring agent such as peppermint, oil of wintergreen, or cherry flavoring may be added. When the unit dosage form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier, such as a vegetable oil or a polyethylene glycol. Various other materials may be present as coatings or to otherwise modify the physical form of the solid unit dosage form. For instance, tablets, pills, or capsules may be coated with gelatin, wax, shellac or sugar and the like. A syrup or elixir may contain the active compound, sucrose or fructose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavoring such as cherry or orange flavor. Of course, any material used in preparing any unit dosage form should be pharmaceutically acceptable and substantially non-toxic in the amounts employed. In addition, the active compound may be incorporated into sustained-release preparations and devices.

Antiviral compounds identified by the present invention and their compositions can also be administered intravenously or intraperitoneally by infusion or injection. Solutions of the active compound or its salts can be prepared in water, optionally mixed with a nontoxic surfactant. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, triacetin, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

Pharmaceutical dosage forms suitable for injection or infusion can include sterile aqueous solutions or dispersions or sterile powders comprising the active ingredient which

are adapted for the extemporaneous preparation of sterile injectable or infusible solutions or dispersions, optionally encapsulated in liposomes. In all cases, the ultimate dosage form should be sterile, fluid and stable under the conditions of manufacture and storage. The liquid carrier or vehicle can be a solvent or liquid dispersion medium comprising, for example, water, ethanol, a polyol (for example, glycerol, propylene glycol, liquid polyethylene glycols, and the like), vegetable oils, nontoxic glyceryl esters, and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the formation of liposomes, by the maintenance of the required particle size in the case of dispersions or by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, buffers or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filter sterilization. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze drying techniques, which yield a powder of the active ingredient plus any additional desired ingredient present in the previously sterile-filtered solutions.

For topical administration, the present compounds may be applied in pure form, i.e., when they are liquids. However, it will generally be desirable to administer them to the skin as compositions or formulations, in combination with a dermatologically acceptable carrier, which may be a solid or a liquid.

Useful solid carriers include finely divided solids such as talc, clay, microcrystalline cellulose, silica, alumina and the like. Useful liquid carriers include water, alcohols or glycols or water-alcohol/glycol blends, in which the present compounds can be dissolved or dispersed at effective levels, optionally with the aid of non-toxic surfactants. Adjuvants such as fragrances and additional antimicrobial agents can be added to optimize the properties for a given use. The resultant liquid compositions can be applied from absorbent pads, used to impregnate bandages and other dressings, or sprayed onto the affected area using pump-type or aerosol sprayers. Thickeners such as synthetic polymers, fatty acids,

fatty acid salts and esters, fatty alcohols, modified celluloses or modified mineral materials can also be employed with liquid carriers to form spreadable pastes, gels, ointments, soaps, and the like, for application directly to the skin of the user.

Examples of useful dermatological compositions which can be used to deliver the  
5 compounds of formula I to the skin are known to the art; for example, see Jacquet et al. (U.S. Pat. No. 4,608,392), Geria (U.S. Pat. No. 4,992,478), Smith et al. (U.S. Pat. No. 4,559,157) and Wortzman (U.S. Pat. No. 4,820,508).

Useful dosages of the compounds of formula I can be determined by comparing  
their *in vitro* activity, and *in vivo* activity in animal models. Methods for the extrapolation  
10 of effective dosages in mice, and other animals, to humans are known to the art; for example, see U.S. Pat. No. 4,938,949.

The compound is conveniently administered in unit dosage form; for example,  
containing 5 to 1000 mg, conveniently 10 to 750 mg, most conveniently, 50 to 500 mg of  
active ingredient per unit dosage form. The desired dose may conveniently be presented in  
15 a single dose or as divided doses administered at appropriate intervals, for example, as two, three, four or more sub-doses per day. The sub-dose itself may be further divided, e.g., into a number of discrete loosely spaced administrations; such as multiple inhalations from an insufflator or by application of a plurality of drops into the eye.

For internal infections, the compositions can be administered orally or parenterally  
20 at dose levels, calculated as the free base, of about 0.1 to 300 mg/kg, preferably 1.0 to 30 mg/kg of mammal body weight, and can be used in man in a unit dosage form, administered one to four times daily in the amount of 1 to 1000 mg per unit dose.

For parenteral administration or for administration as drops, as for eye infections,  
the compounds are presented in aqueous solution in a concentration of from about 0.1 to  
25 about 10%, more preferably about 0.1 to about 7%. The solution may contain other ingredients, such as emulsifiers, antioxidants or buffers.

Generally, the concentration of the compound(s) of formula I in a liquid  
composition, such as a lotion, will be from about 0.1-25 wt-%, preferably from about 0.5-10  
wt-%. The concentration in a semi-solid or solid composition such as a gel or a powder  
30 will be about 0.1-5 wt-%, preferably about 0.5-2.5 wt-%.

The exact regimen for administration of the compounds and compositions disclosed herein will necessarily be dependent upon the needs of the individual subject being treated, the type of treatment and, of course, the judgment of the attending practitioner.

The antiviral activity of a compound of the invention can be determined using pharmacological models which are well known to the art, or using Test A described below.

The compounds of formula (I) and pharmaceutically acceptable salts thereof are useful as antiviral agents. Thus, they are useful to combat viral infections in animals,  
5 including man. The compounds are generally active against herpes viruses, and are particularly useful against the varicella zoster virus, the Epstein-Barr virus, the herpes simplex virus, the human herpes virus type 8 (HHV-8) and the cytomegalovirus (CMV).

10

## CLAIMS

We claim:

1. A method of selecting compounds that inhibit herpes viruses comprising:
  - a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type herpes virus,
  - 5 b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant herpes virus which is the same strain as the wild type herpes virus,
  - c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
  - d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.
- 10 2. A method of selecting compounds that inhibit herpes viruses comprising:
  - a) measuring  $IC_{50}$  of a compound of interest that inhibits a binding domain mutant herpes virus,
  - b) measuring  $IC_{50}$  of the same compound that inhibits a wild type herpes virus which is
  - 15 the same strain as the mutant herpes virus,
  - c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
  - d) selecting the compound of interest wherein the  $IC_{50}$  of step a is at least 3 times greater than the  $IC_{50}$  of step b.
- 20 3. The method of claim 1 or 2 wherein the herpes virus is HSV-1, HSV-2, HCMV, VZV, EBV, or HHV-8.
4. A method of selecting compounds that inhibit herpes viruses comprising:
  - a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HSV-1,
  - 25 b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HSV-1 which is the same strain as the wild type herpes virus,
  - c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
  - d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.
- 30 5. A method of selecting compounds that inhibit herpes viruses comprising:
  - a) measuring  $IC_{50}$  of a compound of interest that inhibits a binding domain mutant HSV-1,

- b) measuring  $IC_{50}$  of the same compound that inhibits a wild type herpes virus which is the same strain as the mutant HSV-1,  
c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and  
d) selecting the compound of interest wherein the  $IC_{50}$  of step a is at least 3 times  
5 greater than the  $IC_{50}$  of step b.
6. The method of claim 4 or 5 wherein HSV-1 is HSV-1 KOS, HSV-1 F, HSV-1 DJL or HSV-1 Patton.
- 10 7. The method of claim 5 or 6 wherein the mutation of a wild type herpes virus to mutant herpes virus is at amino acid 823 from valine to alanine.
8. A method of selecting compounds that inhibit herpes viruses comprising:  
a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HSV-2,  
15 b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HSV-2 which is the same strain as the wild type herpes virus,  
c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and  
d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.
- 20 9. A method of selecting compounds that inhibit herpes viruses comprising:  
a) measuring  $IC_{50}$  of a compound of interest that inhibits a binding domain mutant HSV-2,  
b) measuring  $IC_{50}$  of the same compound that inhibits a wild type herpes virus which is  
25 the same strain as the mutant HSV-2,  
c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and  
d) selecting the compound of interest wherein the  $IC_{50}$  of step a is at least 3 times greater than the  $IC_{50}$  of step b.
- 30 10. The method of claim 8 or 9 wherein HSV-2 is HSV-2 MS, HSV-2 35D, or HSV-2 186.
11. A method of selecting compounds that inhibit herpes viruses comprising:

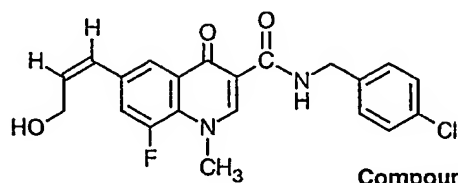
- a) measuring  $IC_{50}$  of a compound of interest that inhibits a wild type HCMV,
  - b) measuring  $IC_{50}$  of the same compound that inhibits a binding domain mutant HCMV which is the same strain as the wild type herpes virus,
  - c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
  - 5 d) selecting the compound of interest wherein the  $IC_{50}$  of step b is at least 3 times greater than the  $IC_{50}$  of step a.
- 
12. A method of selecting compounds that inhibit herpes viruses comprising:
    - a) measuring  $IC_{50}$  of a compound of interest that inhibits a binding domain mutant  
10 HCMV,
    - b) measuring  $IC_{50}$  of the same compound that inhibits a wild type herpes virus which is the same strain of the mutant HCMV,
    - c) comparing  $IC_{50}$  of step a with  $IC_{50}$  of step b; and
    - d) selecting the compound of interest wherein the  $IC_{50}$  of step a is at least 3 times  
15 greater than the  $IC_{50}$  of step b.
- 
13. The method of claim 8 or 9 wherein HCMV is AD169.
- 
14. The methods of claims 1, 4, 8, or 11 wherein  $IC_{50}$  of step b is at least 5 times greater  
20 than the  $IC_{50}$  of step a.
- 
15. The methods of claims 2, 5, 9, or 12 wherein  $IC_{50}$  of step a is at least 5 times greater than the  $IC_{50}$  of step b.
- 
- 25 16. A use of compounds for manufacturing of medicinals for selectively treating diseases caused by herpes viruses in a human host comprising administering a compound to a human in need of such treatment wherein said compound inhibits herpes viruses by interaction with the binding domain in the viral DNA polymerase.
- 
- 30 17. A use of compounds for manufacturing of medicinals for selectively inhibiting herpes viruses in a human host comprising administering a compound to a human in need of such treatment wherein  $IC_{50}$  of the compound that inhibits a binding domain

mutant herpes virus is at least 3 times greater than  $IC_{50}$  of the compound that inhibits a wild type herpes virus which is the same strain as the mutant herpes virus.

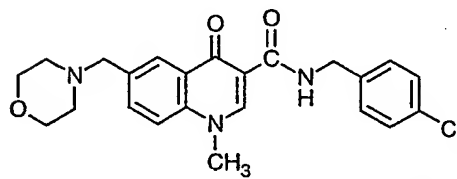
18. The use of claim 17 wherein  $IC_{50}$  of the compound that inhibits a binding domain mutant herpes virus is at least 5 times greater than  $IC_{50}$  of the compound that inhibits a wild type herpes virus which is the same strain as the mutant herpes virus.
19. The use of claim 17 wherein herpes viruses is HSV-1, HSV-2, HCMV, VZV, EBV, or HHV-8.
20. A use of compounds for manufacturing of medicinals for treating herpesviral infections in a human host wherein  $IC_{50}$  of the compound that inhibits a binding domain mutant herpes virus is at least 5 times greater than  $IC_{50}$  of the compound that inhibits a wild type herpes virus which is the same strain as the mutant herpes virus.
21. A use of compounds for manufacturing of medicinals for treating herpesviral infections in a human host wherein said compound inhibits the herpesvirus by interacting with the binding domain in the viral DNA polymerase.
22. The herpesviral infection of claim 20 or 21 which is HSV-1, HSV-2, HCMV, VZV, EBV, or HHV-8 infection.
23. A compound for the inhibiting of herpesvirus DNA polymerases wherein passage of a wild type herpes virus in the presence of said compound results a change of the wild type HSV-1 polymerases at amino acid 823 from valine to alanine.
24. A compound for inhibiting herpesvirus DNA polymerases wherein passage of a wild type herpes virus in the presence of said compound results in a change of the wild type HCMV polymerases at amino acid 823 from valine to alanine and at amino acid 824 from valine to leucine.

25. A mutant herpesvirus DNA molecule having a nucleotide sequence selected from a group consisting of SEQ.ID.NO. 1; SEQ.ID.NO. 3; SEQ.ID.NO. 5; SEQ.ID.NO. 7; SEQ.ID.NO. 9; and SEQ.ID.NO. 11.
- 5 26. A mutant herpesvirus polymerase amino acid molecule having an amino acid sequence selected from a group consisting of SEQ.ID.NO. 2; SEQ.ID.NO. 4; SEQ.ID.NO. 6; SEQ.ID.NO. 8; SEQ.ID.NO. 10 and SEQ.ID.NO. 12.

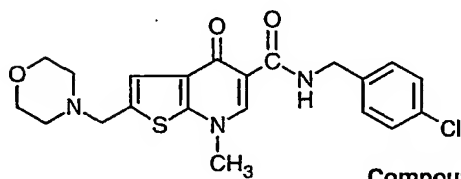
**Figure 1** 4-HQ, 4-oxo-DHQ and 4-oxo-DHTP antiviral compounds



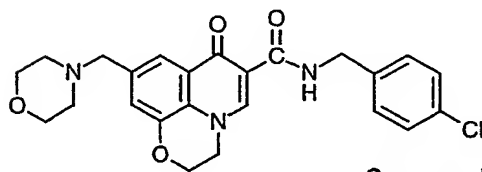
**Compound No. 1**



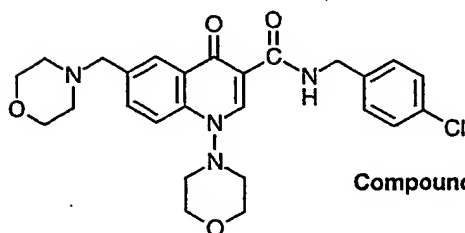
**Compound No. 2**



**Compound No. 3**

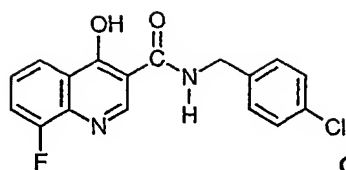


**Compound No. 4**

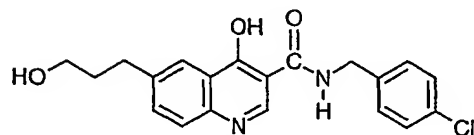


**Compound No. 5**

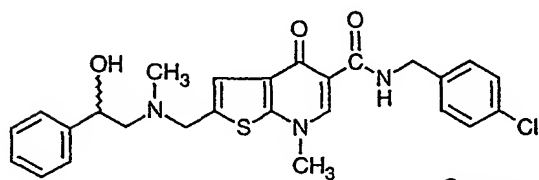
(Figure 1 continue)



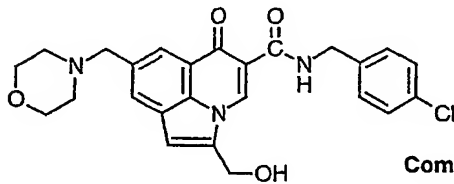
Compound No. 6



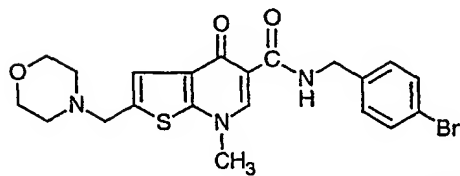
Compound No. 7



Compound No. 8

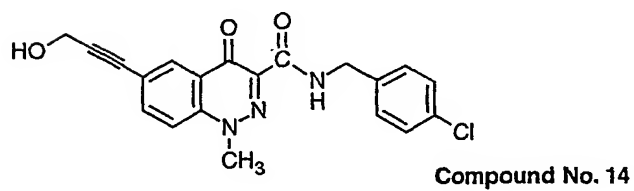
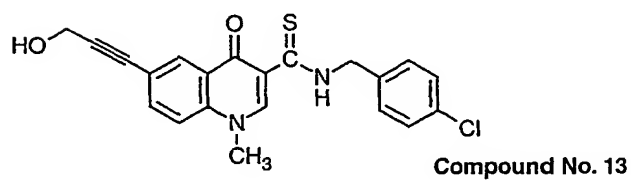
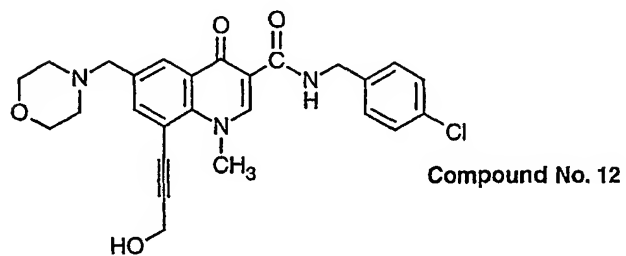
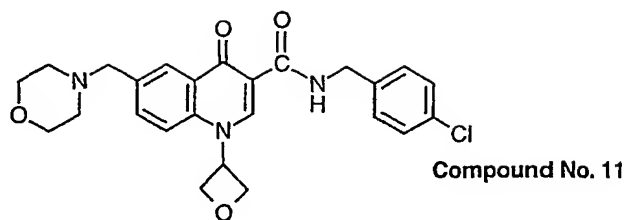


Compound No. 9

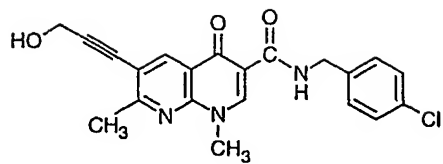


Compound No. 10

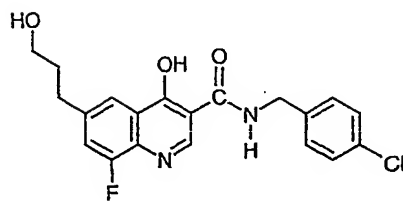
(Figure 1 continue)



(Figure 1 continue)

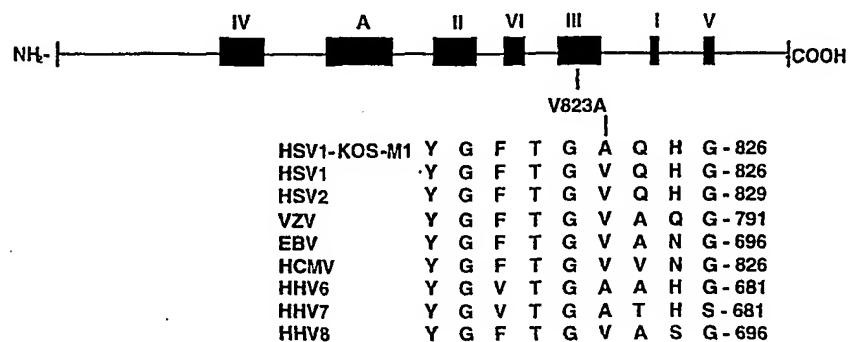


Compound No.15

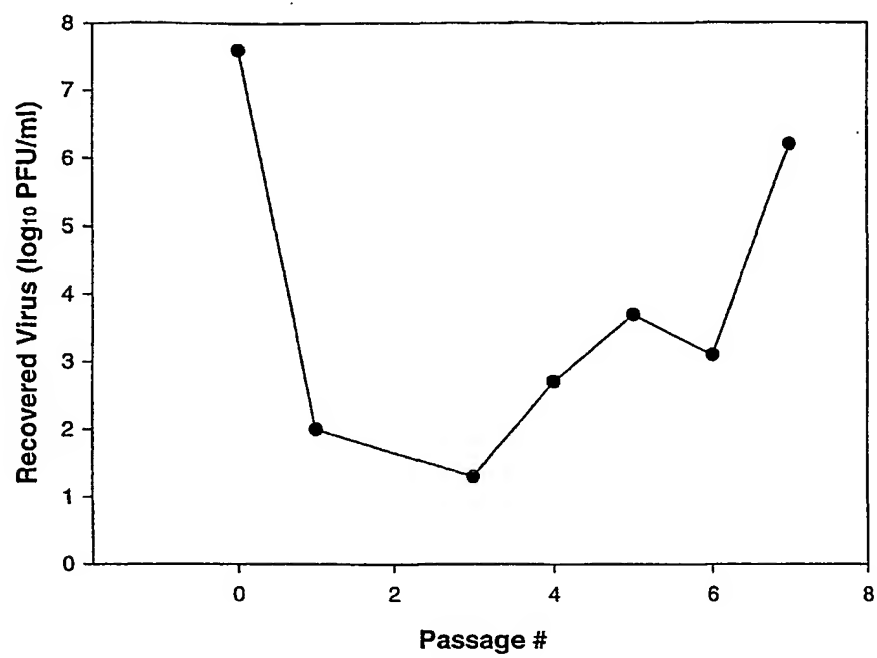


Compound 17

**Figure 2. The HSV1 (KOS Strain) DNA Polymerase Amino Acid 823 is Critical for Resistance to 4-Hydroxyquinolines and Related Compounds**



Schematic of HSV1 polymerase illustrating the conserved regions A and I-VI found in class 2 polymerases. Also shown are the amino acid sequence for the highly conserved herpesvirus domain in region III which surrounds the HSV1 amino acid 823.

**Figure 3** Serial Passage of HSV-1 in Presence of 20  $\mu$ M compound 17

**Figure 4 Comparison of Wild type HSV-1 and HSV-2 DNA Polymerases Amino Acid Sequences Alligned by Amino Acid Homology\***

5	HSV2-MS	MFCAAGGPTS	PGGKSAARAA	SGFFAPHNPR	GATQTAPPPC	RRQNFYNPHL	-50
	HSV2-186	MFCAAGGPAS	PGGKSAARAA	SGFFAPHNPR	GATQTAPPPC	RRQNFYNPHL	-50
	HSV1-Kos	MFSGGGGPLS	PGGKSAARAA	SGFFAPAGPR	GAGR.GPPPC	LRQNFYNPYL	-49
	HSV1-Patton	MFSGGGGPLS	PGGKSAARAA	SGFFAPAGPR	GAGR.GPPPC	LRQNFYNPYL	-49
	HSV1-DJL	MFSGGGGPLS	PGGKSAARAA	SGFFAPAGPR	GAGR.GPPPC	LRQNFYNPYL	-49
10	HSV1-F	MFSGGGGPLS	PGGKSAARAA	SGFFAPAGPR	GAGR.GPPPC	LRQNFYNPYL	-49
	HSV2-MS	AQTGTQPKAP	GPAQRHTYYS	ECDEFRFIAP	RSLDEDAPAE	QRTGVHDGRL	-100
	HSV2-186	AQTGTQPKAP	GPAQRHTYYS	ECDEFRFIAP	RSLDEDAPAE	QRTGVHDGRL	-100
	HSV1-Kos	APVGTQQKPT	GPTQRHTYYS	ECDEFRFIAP	RVLDEDAPPE	KRAGVHDGHL	-99
	HSV1-Patton	APVGTQQKPT	GPTQRHTYYS	ECDEFRFIAP	RVLDEDAPPE	KRAGVHDGHL	-99
15	HSV1-DJL	APVGTQQKPT	GPTQRHTYYS	ECDEFRFIAP	RVLDEDAPPE	KRAGVHDGHL	-99
	HSV1-F	APVGTQQKPT	GPTQRHTYYS	ECDEFRFIAP	RVLDEDAPPE	KRAGVHDGHL	-99
20	HSV2-MS	RRAPKVYCGG	DERDVLRVGP	EGFWPRRLRL	WGGADHAPKG	FDPTVTVFHV	-150
	HSV2-186	RRAPKVYCGG	DERDVLRVGP	EGFWPRRLRL	WGGADHAPKG	FDPTVTVFHV	-150
	HSV-Kos	KRAPKVYCGG	DERDVLRVGS	GGFWPRRSRL	WGGVDHAPAG	FNPTVTVFHV	-149
	HSV1-Patton	KRAPKVYCGG	DERDVLRVGS	GGFWPRRSRL	WGGVDHAPAG	FNPTVTVFHV	-149
	HSV1-DJL	KRAPKVYCGG	DERDVLRVGS	GGFWPRRSRL	WGGVDHAPAG	FNPTVTVFHV	-149
25	HSV1-F	KRAPKVYCGG	DERDVLRVGS	GGFWPRRSRL	WGGVDHAPAG	FNPTVTVFHV	-149
	HSV2-MS	YDILEHVEHA	YSMRAAQLHE	RFMDAITPAG	TVITLLGLTP	EGHRVAVHVY	-200
	HSV2-186	YDILEHVEHA	YSMRAAQLHE	RFMDAITPAG	TVITLLGLTP	EGHRVAVHVY	-200
	HSV-Kos	YDILENVEHA	YGMRAAQFHA	RFMDAITPTG	TVITLLGLTP	EGHRVAVHVY	-199
	HSV1-Patton	YDILENVEHA	YGMRAAQFHA	RFMDAITPTG	TVITLLGLTP	EGHRVAVHVY	-199
30	HSV1-DJL	YDILENVEHA	YGMRAAQFHA	RFMDAITPTG	TVITLLGLTP	EGHRVAVHVY	-199
	HSV1-F	YDILENVEHA	YGMRAAQFHA	RFMDAITPTG	TVITLLGLTP	EGHRVAVHVY	-199
35	HSV2-MS	GTRQYFYMNK	AEVDRHLQCR	APRDL CERLA	AALRESPGAS	FRGISADHFE	-250
	HSV2-186	GTRQYFYMNK	AEVDRHLQCR	APRDL CERLA	AALRESPGAS	FRGISADHFE	-250
	HSV-Kos	GTRQYFYMNK	EEVDRHLQCR	APRDL CERMA	AALRESPGAS	FRGISADHFE	-249
	HSV1-Patton	GTRQYFYMNK	EEVDRHLQCR	APRDL CERMA	AALRESPGAS	FRGISADHFE	-249
	HSV1-DJL	GTRQYFYMNK	EEVDRHLQCR	APRDL CERMA	AALRESPGAS	FRGISADHFE	-249
40	HSV1-F	GTRQYFYMNK	EEVDRHLQCR	APRDL CERMA	AALRESPGAS	FRGISADHFE	-249
	HSV2-MS	AEVVERADVY	YYETRPTLYY	RVFVRSGRAL	AYLCDNF CPA	IRKYE GGVDA	-300
	HSV2-186	AEVVERADVY	YYETRPTLYY	RVFVRSGRAL	AYLCDNF CPA	IRKYE GGVDA	-300
	HSV-Kos	AEVVERTDVY	YYETRPALFY	RVYVRSGRVL	SYLCDNF CPA	IKKYE GGVDA	-299
	HSV1-Patton	AEVVERTDVY	YYETRPALFY	RVYVRSGRVL	SYLCDNF CPA	IKKYE GGVDA	-299
45	HSV1-DJL	AEVVERTDVY	YYETRPALFY	RVYVRSGRVL	SYLCDNF CPA	IKKYE GGVDA	-299
	HSV1-F	AEVVERTDVY	YYETRPALFY	RVYVRSGRVL	SYLCDNF CPA	IKKYE GGVDA	-299
50	HSV2-MS	TTRFILDNPG	FVTFGWYRLK	PGRGNAPAQP	RPPTAFGTSS	DVEFNCTADN	-350
	HSV2-186	TTRFILDNPG	FVTFGWYRLK	PGRGNAPAQP	RPPTAFGTSS	DVEFNCTADN	-350
	HSV-Kos	TTRFILDNPG	FVTFGWYRLK	PGRNNTLAQP	RAPMAFGTSS	DVEFNCTADN	-349
	HSV1-Patton	TTRFILDNPG	FVTFGWYRLK	PGRNNTLAQP	RAPMAFGTSS	DVEFNCTADN	-349
	HSV1-DJL	TTRFILDNPG	FVTFGWYRLK	PGRNNTLAQP	RAPMAFGTSS	DVEFNCTADN	-349
55	HSV1-F	TTRFILDNPG	FVTFGWYRLK	PGRNNTLAQP	RAPMAFGTSS	DVEFNCTADN	-349
	HSV2-MS	LAVEGAMCDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAERPED	LVIQISCLLY	-400
	HSV2-186	LAVEGAMCDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAERPED	LVIQISCLLY	-400
	HSV-Kos	LAIEGGMSDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAGHPED	LVIQISCLLY	-399
	HSV1-Patton	LAIEGGMSDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAGHPED	LVIQISCLLY	-399
60	HSV1-DJL	LAIEGGMSDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAGHPED	LVIQISCLLY	-399
	HSV1-F	LAIEGGMSDL	PAYKLMCFDI	ECKAGGEDEL	AFPVAGHPED	LVIQISCLLY	-399
65	HSV2-MS	DLSTTALEHI	LLFSLGSCDL	PESHLSDLAS	RGLPAPVVLE	FDSEFEMLLA	-450
	HSV2-186	DLSTTALEHI	LLFSLGSCDL	PESHLSDLAS	RGLPAPVVLE	FDSEFEMLLA	-450
	HSV-Kos	DLSTTALEHV	LLFSLGSCDL	PESHLNELAA	RGLPTPVVLE	FDSEFEMLLA	-449
	HSV1-Patton	DLSTTALEHV	LLFSLGSCDL	PESHLNELAA	RGLPTPVVLE	FDSEFEMLLA	-449
	HSV1-DJL	DLSTTALEHV	LLFSLGSCDL	PESHLNELAA	RGLPTPVVLE	FDSEFEMLLA	-449
65	HSV1-F	DLSTTALEHV	LLFSLGSCDL	PESHLNELAA	RGLPTPVVLE	FDSEFEMLLA	-449

5	HSV2-MS	FMTFVKQYGP	EFVTGYNIIIN	FDWPFVLTKL	TEIYKVPLDG	YGRMNGRGVF	-500
	HSV2-186	FMTFVKQYGP	EFVTGYNIIIN	FDWPFVLTKL	TEIYKVPLDG	YGRMNGRGVF	-500
	HSV-Kos	FMTLVKQYGP	EFVTGYNIIIN	FDWPFLLAKL	TDIYKVPLDG	YGRMNGRGVF	-499
	HSV1-Patton	FMTLVKQYGP	EFVTGYNIIIN	FDWPFLLAKL	TDIYKVPLDG	YGRMNGRGVF	-499
	HSV1-DJL	FMTLVKQYGP	EFVTGYNIIIN	FDWPFLLAKL	TDIYKVPLDG	YGRMNGRGVF	-499
10	HSV1-F	FMTLVKQYGP	EFVTGYNIIIN	FDWPFLLAKL	TDIYKVPLDG	YGRMNGRGVF	-499
	HSV2-MS	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKVKLSSYK	LNAVAEAVLK	-550
	HSV2-186	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKVKLSSYK	LNAVAEAVLK	-550
	HSV-Kos	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKIKLSSYK	LNAVAEAVLK	-549
	HSV1-Patton	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKIKLSSYK	LNAVAEAVLK	-549
15	HSV1-DJL	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKIKLSSYK	LNAVAEAVLK	-549
	HSV1-F	RVWDIGQSHF	QKRISKIKVNG	MVNIDMYGII	TDKIKLSSYK	LNAVAEAVLK	-549
	HSV2-MS	DKKKDLSYRD	IPAYYASGPA	QRGVIGEYCV	QDSLLVGQLF	FKFLPHLELS	-600
	HSV2-186	DKKKDLSYRD	IPAYYASGPA	QRGVIGEYCV	QDSLLVGQLF	FKFLPHLELS	-600
	HSV-Kos	DKKKDLSYRD	IPAYYAAGPA	QRGVIGEYCI	QDSLLVGQLF	FKFLPHLELS	-599
20	HSV1-Patton	DKKKDLSYRD	IPAYYAAGPA	QRGVIGEYCI	QDSLLVGQLF	FKFLPHLELS	-599
	HSV1-DJL	DKKKDLSYRD	IPAYYAAGPA	QRGVIGEYCI	QDSLLVGQLF	FKFLPHLELS	-599
	HSV1-F	DKKKDLSYRD	IPAYYAAGPA	QRGVIGEYCI	QDSLLVGQLF	FKFLPHLELS	-599
	HSV2-MS	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	GQKGFILPDT	QGRFRGLDKE	-650
	HSV2-186	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	GQKGFILPDT	QGRFRGLDKE	-650
25	HSV-Kos	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	DQKGFILPDT	QGRFRGAGGE	-649
	HSV1-Patton	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	DQKGFILPDT	QGRFRGAGGE	-649
	HSV1-DJL	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	DQKGFILPDT	QGRFRGAGGE	-649
	HSV1-F	AVARLAGINI	TRTIYDGQOI	RVFTCLLRLA	DQKGFILPDT	QGRFRGGGGE	-649
30	HSV2-MS	APKRPAVPRG	EGERP GDGNG	DEDKDDDE..	DEDGDERE.E	VARETGGRHV	-697
	HSV2-186	APKRPAVPRG	EGERP GDGNG	DEDKDDDEDG	DEDGDERE.E	VARETGGRHV	-697
	HSV-Kos	APKRPAARE	DEERP.....	EEGEDEDER	EEGGGEREPE	GARETAGRHV	-694
	HSV1-Patton	APKRPAARE	DEERP.....	EEGEDEDER	EEGGGEREPE	GARETAGRHV	-694
	HSV1-DJL	APKRPAARE	DEERP.....	EEGEDENER	EEGGGEREPE	GARETAGRHV	-694
35	HSV1-F	APKRPAARE	DEERP.....	EEGEDEDER	EEGGGEREPE	GARETAGRHV	-694
	HSV2-MS	GYQGARVLDP	TSGFHVDPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRPEAVA	-747
	HSV2-186	GYQGARVLDP	TSGFHVDPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRPEAVA	-749
	HSV-Kos	GYQGARVLDP	TSGFHVNPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRADAVA	-744
	HSV1-Patton	GYQGARVLDP	ISGFHVNPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRADAVA	-744
40	HSV1-DJL	GYQGARVLDP	TSGFHVNPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRADAVA	-744
	HSV1-F	GYQGARVLDP	TSGFHVNPVV	VDFDASLYPS	IIQAHNLCFS	TLSLRADAVA	-744
	HSV2-MS	HLEADRDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-797
	HSV2-186	HLEADRDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-799
	HSV-Kos	HLEAGKDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-794
45	HSV1-Patton	HLEAGKDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-794
	HSV1-DJL	HLEAGKDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-794
	HSV1-F	HLEAGKDYLE	IEVGGRRLFF	VKAHVRESLL	SILLRDWLAM	RKQIRSRIPQ	-794
50	HSV2-MS	STPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-847
	HSV2-186	SPPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-849
	HSV-Kos	SSPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-844
	HSV1-Patton	SSPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-844
	HSV1-DJL	SSPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-844
55	HSV1-F	SSPEEAVLLD	KQQAIAKVVC	NSVYGFTGVQ	HGLLPCLHVA	ATVTTIGREM	-844
	HSV2-MS	LLATRAYVHA	RWAEDQLLA	DFPEAAGMRA	PGPYSMRIIY	GDTDSIFVLC	-897
	HSV2-186	LLATRAYVHA	RWAEDQLLA	DFPEAAGMRA	PGPYSMRIIY	GDTDSIFVLC	-899
	HSV-Kos	LLATREYVHA	RWAAFEQLLA	DFPEAADMRA	PGPYSMRIIY	GDTDSIFVLC	-894
	HSV1-Patton	LLATREYVHA	RWAAFEQLLA	DFPEAADMRA	PGPYSMRIIY	GDTDSIFVLC	-894
60	HSV1-DJL	LLATREYVHA	RWAAFEQLLA	DFPEAADMRA	PGPYSMRIIY	GDTDSIFVLC	-894
	HSV1-F	LLATREYVHA	RWAAFEQLLA	DFPEAADMRA	PGPYSMRIIY	GDTDSIFVLC	-894
	HSV2-MS	RGLTAAGLVA	MGDKMASHIS	RALFLPPIKL	ECEKTFTKLL	LIAKKYYIGV	-947
	HSV2-186	RGLTAAGLVA	MGDKMASHIS	RALFLPPIKL	ECEKTFTKLL	LIAKKYYIGV	-949
	HSV-Kos	RGLTAAGLTA	MGDKMASHIS	RALFLPPIKL	ECEKTFTKLL	LIAKKYYIGV	-944
65	HSV1-Patton	RGLTAAGLTA	MGDKMASHIS	RALFLPPIKL	ECEKTFTKLL	LIAKKYYIGV	-944

	HSV1-DJL	RGHTAAGLTA	VGDKMASHIS	RALFLPPIKL	ECEKTFTKLL	LIAKKKYIGV	-944
	HSV1-F	RGHTAAGLTA	VGDKMASHIS	RALFLSPIKL	ECEKTFTKLL	LIAKKKYIGV	-944
	HSV2-MS	ICGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-997
5	HSV2-186	ICGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-999
	HSV-Kos	IYGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-994
	HSV1-Patton	IYGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-994
	HSV1-DJL	IYGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-994
	HSV1-F	IYGGKMLIKG	VDLVRKNNCA	FINRTSRALV	DLLFYDDTVS	GAAAAAERP	-994
10	HSV2-MS	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1047
	HSV2-186	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1049
	HSV-Kos	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1044
	HSV1-Patton	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1044
15	HSV1-DJL	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1044
	HSV1-F	AEEWLARPLP	EGLQAFGAVL	VDHRRITDP	ERDIQDFVLT	AELSRHPRAY	-1044
	HSV2-MS	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1097
	HSV2-186	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1099
20	HSV-Kos	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1094
	HSV1-Patton	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1094
	HSV1-DJL	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1094
	HSV1-F	TNKRLAHLTV	YYKLARRAQ	VPSIKDRIPY	VIVAQTREVE	ETVARLAALR	-1094
25	HSV2-MS	ELDAAAPGDE	PAPPAALPSP	AKRPRETPSH	ADPPGGASKP	RKLLVSELAE	-1147
	HSV2-186	ELDAAAPGDE	PAPPAALPSP	AKRPRETPSH	ADPPGGASKP	RKLLVSELAE	-1149
	HSV-Kos	ELDAAAPGDE	PAPPAALPSP	AKRPRETPSH	ADPPGGASKP	RKLLVSELAE	-1144
	HSV1-Patton	ELDAAAPGDE	PAPPAALPSP	AKRPRETPSP	ADPPGGASKP	RKLLVSELAE	-1144
	HSV1-DJL	ELDAAAPGDE	PAPPAALPSP	AKRPRETPSP	ADPPGGASKP	RKLLVSELAE	-1144
30	HSV1-F	ELDAAAPGDE	PAPPAALPSP	AKRPRETPLH	ADPPGGASKP	RKLLVSELAE	-1144
	HSV2-MS	DPGYAARGV	PLNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1197
	HSV2-186	DPGYAARGV	PLNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1199
	HSV-Kos	DPAYAIAHGV	ALNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1194
35	HSV1-Patton	DPAYAIAHGV	ALNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1194
	HSV1-DJL	DPAYAIAHGV	ALNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1194
	HSV1-F	DPAYAIAHGV	ALNTDYYFSH	LLGAACVTFK	ALFGNNAKIT	ESLLKRFIPE	-1194
	HSV2-MS	TWHPPDDVAA	RLRAAGFGPA	GAGATAEETR	RMLHRAFDL	A* -1238	
40	HSV2-186	TWHPPDDVAA	RLRAAGFGPA	GAGATAEETR	RMLHRAFDL	A* -1240	
	HSV-Kos	VWHPPDDVAA	RLRAAGFGAV	GAGATAEETR	RMLHRAFDL	A* -1235	
	HSV1-Patton	VWHPPDDVTA	RLRAAGFGAV	GAGATAEETR	RMLHRAFDL	A* -1235	
	HSV1-DJL	VWHPPDDVAA	RLRTAGFGAV	GAGATAEETR	RMLHRAFDL	A* -1235	
	HSV1-F	VWHPPDDVAA	RLRAAGFGAV	GAGATAEETR	RMLHRAFDL	A* -1235	

45

\*Amino acid alignment demonstrates difference in amino acid's sequences.

\*The gaps "....." indicate missing amino acids relative to other stanins.

\*Wild HSV2-MS is listed as SEQ. ID NO 14.

\*Wild HSV2-186 is listed as SEQ. ID NO 15.

50 \*Wild HSV-Kos is listed as SEQ. ID NO 16.

\*Wild HSV1-Patton is listed as SEQ. ID NO 17.

\*Wild HSV1-DJL is listed as SEQ. ID NO 18.

\*Wild HSV1-F is listed as SEQ. ID NO 19.

55

**Figure 5** DNA and amino acid sequence list

**SEQ. ID. NO. 1** DNA sequence of DNA polymerase gene for HSV2-MS-M1

```

5   1 ATGTTTGTG CCGCGGGCGG CCCGACTTCC CCCGGGGGGA AGTCGGCGGC
    51 TCGGGCGGCG TCTGGGTTTT TTGCCCCCA CAACCCCCGG GGAGCCACCC
    101 AGACGGCACC GCCGCCTTGC CGCCGGCAGA ACTTCTACAA CCCCACCTC
10  151 GCTCAGACCG GAACGCAGCC AAAGGCCCCC GGGCCGGCTC AGCGCCATAC
    201 GTACTACAGC GAGTGCGACG AATTTCGATT TATCGCCCCG CGTTCGCTGG
    251 ACGAGGACGC CCCC GCGGAG CAGCGCACCG GGGTCCACGA CGGCCGCCTC
    301 CGGCGCGCCC CTAAGGTGTA CTGCGGGGGG GACGAGCGCG ACGTCCTCCG
    351 CGTGGGCCCC GAGGGCTTCT GGCCGCGTCG CTTGCGCCTG TGGGGCGGTG
20  401 CGGACCATGC CCCCAAGGGG TTCGACCCCA CCGTCACCGT CTTCACGTG
    451 TACGACATCC TGGAGCACGT GGAACACGCG TACAGCATGC GCGCCGCCCA
    501 GCTCCACGAG CGATTTATGG ACGCCATCAC GCCCGCCGGG ACCGTCATCA
25  551 CGCTTCTGGG TCTGACCCCC GAAGGCCATC GCGTCGCCGT TCACGTCTAC
    601 GGCACGCGGC AGTACTTTTA CATGAACAAG GCGGAGGTGG ATCGGCACCT
30  651 GCAGTGCCGT GCCCCGCGCG ATCTCTGCGA GCGCCTGGCG GCGGCCCTGC
    701 GCGAGTCGCC GGGGGCGTCG TTCCGCGGCA TCTCCGCGGA CCACTTCGAG
    751 GCGGAGGTGG TGGAGCGCGC CGACGTGTAC TATTACGAAA CGCGCCCAGC
35  801 CCTGTACTAC CGCGTCTTCG TGCGAAGCGG GCGCGCGCTG GCCTACCTGT
    851 GCGACAACTT TTGCCCCGCG ATCAGGAAGT ACGAGGGGGG CGTCGACGCC
40  901 ACCACCCGGT TTATCCTGGA CAACCCGGGG TTTGTACCT TCGGCTGGTA
    951 CCGCCTCAAG CCCGGCCGCG GGAACGCGCC GGCCCAACCG CGCCCCCGA
45  1001 CGGCGTTCGG AACCTCGAGC GACGTCGAGT TTAAGTGCAC GGCGGACAAC
    1051 CTGGCCGTCG AGGGGGCCAT GTGTGACCTG CCGGCCTACA AGCTCATGTG
    1101 CTTCGATATC GAATGCAAGG CCGGGGGGGA GGACGAGCTG GCCTTCCGG
50  1151 TCGCGGAACG CCCGGAAGAC CTCGTCATCC AGATCTCCTG TCTGCTCTAC
    1201 GACCTGTCCA CCACGCCCT CGAGCACATC CTCCTGTTTT CGCTCGGATC
55  1251 CTGCGACCTC CCCGAGTCCC ACCTCAGCGA TCTCGCCTCC AGGGGCCTGC
    1301 CGGCCCCCGT CGTCCTGGAG TTTGACAGCG AATTCGAGAT GCTGCTGGCC

```

1351 TTCATGACCT TCGTCAAGCA GTACGGCCCC GAGTTCGTGA CCGGGTACAA  
1401 CATCATCAAC TTCGACTGGC CCTTCGTCCT GACCAAGCTG ACGGAGATCT  
5 1451 ACAAGGTCCC GCTCGACGGG TACGGGCGCA TGAACGGCCG GGGTGTGTTC  
1501 CGCGTGTGGG ACATCGGCCA GAGCCACTTT CAGAAGCGCA GCAAGATCAA  
1551 GGTGAACGGG ATGGTGAACA TCGACATGTA CGGCATCATC ACCGACAAGG  
10 1601 TCAAACCTCTC CAGCTACAAG CTGAACGCCG TCGCCGAGGC CGTCTTGAAG  
1651 GACAAGAAGA AGGATCTGAG CTACCGCGAC ATCCCCGCCT ACTACGCCTC  
1701 CGGGCCCCGCG CAGCGCGGGG TGATCGGCGA GTATTGTGTG CAGGACTCGC  
15 1751 TGCTGGTCGG GCAGCTGTTT TTCAAGTTTC TGCCGCACCT GGAGCTTTCC  
1801 GCCGTCGCGC GCCTGGCGGG CATCAACATC ACCCGCACCA TCTACGACGG  
1851 CCAGCAGATC CGCGTCTTCA CGTGCCTCCT GCGCCTTGCG GGCCAGAAGG  
1901 GCTTCATCCT GCCGGACACC CAGGGGCGGT TTCGGGGCCT CGACAAGGAG  
20 1951 GCGCCCAAGC GCGCGGCCGT GCCTCGGGGG GAAGGGGAGC GGCCGGGGGA  
2001 CGGGAACGGG GACGAGGATA AGGACGACGA CGAGGACGAG GACGGGGACG  
2051 AGCGCGAGGA GGTCGCGCGC GAGACCGGGG GCCGGCACGT TGGGTACCAG  
2101 GGGGCCCCGGG TCCTCGACCC CACCTCCGGG TTTCACGTCG ACCCCGTGGT  
2151 GGTGTTTGAC TTTGCCAGCC TGTACCCAG CATCATCCAG GCCACAACC  
2201 TGTGCTTCAG TACGCTCTCC CTGCGGCCCG AGGCCGTCGC GCACCTGGAG  
2251 GCGGACCGGG ACTACCTGGA GATCGAGGTG GGGGGCCGAC GGCTGTTCTT  
2301 CGTGAAGGCC CACGTACGCG AGAGCCTGCT GAGCATCCTG CTGCGCGACT  
2351 GGCTGGCCAT GCGAAAGCAG ATCCGCTCGC GGATCCCCCA GAGCACCCCC  
2401 GAGGAGGCCG TCCTCCTCGA CAAGCAACAG GCCGCCATCA AGGTGGTGTG  
2451 CAACTCGGTG TACGGGTTC ACGGGGCGCA GCACGGTCTT CTGCCCTGCC  
2501 TGCACGTGGC CGCCACCGTG ACGACCATCG GCCGCGAGAT GCTCCTCGCG  
2551 ACGCGCGCGT ACGTGACGC GCGCTGGGCG GAGTTCGATC AGCTGCTGGC  
2601 CGACTTTCCG GAGGCGGCCG GCATGCGCGC CCCCAGTCCG TACTCCATGC  
2651 GCATCATCTA CGGGGACACG GACTCCATTT TCGTTTTGTG CCGCGGCCTC  
2701 ACGGCCGCGG GCCTGGTGGC CATGGGCGAC AAGATGGCGA GCCACATCTC  
2751 GCGCGCGCTG TTCCTCCCC CGATCAAGCT CGAGTGCGAA AAAACGTTCA  
2801 CCAAGCTGCT GCTCATCGCC AAGAAAAAGT ACATCGGCGT CATCTGCGGG  
60

2851 GGCAAGATGC TCATCAAGGG CGTGGATCTG GTGCGCAAAA ACAACTGCGC  
2901 GTTTATCAAC CGCACCTCCA GGGCCCTGGT CGACCTGCTG TTTTACGACG  
5 2951 ATACCGTATC CGGAGCGGCC GCCGCGTTAG CCGAGCGCCC CGCAGAGGAG  
3001 TGGCTGGCGC GACCCCTGCC CGAGGGACTG CAGGCGTTCG GGGCCGTCCT  
3051 CGTAGACGCC CATCGGCGCA TCACCGACCC GGAGAGGGAC ATCCAGGACT  
10 3101 TTGTCCTCAC CGCCGAAGT AGCAGACACC CGCGCGCGTA CACCAACAAG  
3151 CGCCTGGCCC ACCTGACGGT GTATTACAAG CTCATGGCCC GCCGCGCGCA  
15 3201 GGTCCCGTCC ATCAAGGACC GGATCCCGTA CGTGATCGTG GCCCAGACCC  
3251 GCGAGGTAGA GGAGACGGTC GCGCGGCTGG CCGCCCTCCG CGAGCTAGAC  
3301 GCCGCCGCCC CAGGGGACGA GCCCGCCCCC CCAGCGGCCC TGCCCTCCCC  
20 3351 GGCCAAGCGC CCCCGGGAGA CGCCGTCGCA TGCCGACCCC CCGGGAGGCG  
3401 CGTCCAAGCC CCGCAAGCTG CTGGTGTCCG AGCTGGCGGA GGATCCCGGG  
25 3451 TACGCCATCG CCCGGGGCGT TCCGCTCAAC ACGGACTATT ACTTCTCGCA  
3501 CCTGCTGGGG GCGGCCTGCG TGACGTTCAA GGCCCTGTTT GGAAATAACG  
3551 CCAAGATCAC CGAGAGTCTG TTAAAGAGGT TTATTCCCGA GACGTGGCAC  
30 3601 CCCCCGGACG ACGTGGCCGC GCGGCTCAGG GCCGCGGGGT TCGGGCCGGC  
3651 GGGGGCCGGC GCTACGGCGG AGGAAACTCG TCGAATGTTG CATAGAGCCT  
35 3701 TTGATACTCT AGCATGA

## SEQ. ID. NO. 2      Amino acid sequence of DNA polymerase for HSV2-MS-M1

1 MFCAAGGPTS PGGKSAARAA SGFFAPHNPR GATQTAPPPC RRQNFYNPHL  
5 51 AQTGTQPKAP GPAQRHTYYS ECDEFRIAP RSLDEDAPAE QRTGVHDGRL  
101 RRAPKVYCGG DERDVLRVGP EGFWRRLRL WGGADHAPKG FDPTVTVFHV  
151 YDILEHVEHA YSMRAAQLHE RFMDAITPAG TVITLLGLTP EGHRVAVHVY  
10 201 GTRQYFYMNK AEVDRHLQCR APRDLCERLA AALRESPGAS FRGISADHFE  
251 AEVVERADVY YYETRPTLYY RVFVRSGRAL AYLCDNFCA IRKYEGGVDA  
15 301 TTRFILDNPG FVTFGWYRLK PGRGNAPAQP RPPTAFGTSS DVEFNCTADN  
351 LAVEGAMCDL PAYKLMCFDI ECKAGGEDEL AFPVAERPED LVIQISCLLY  
401 DLSTTALEHI LLFSLGSCDL PESHLSDLAS RGLPAPVVLE FDSEFEMLLA  
20 451 FMTFVKQYGP EFVTGYNIN FDWPFVLTCL TEIYKVPLDG YGRMNGRGVF  
501 RVWDIGQSHF QKRSKIKVNG MVNIDMYGII TDKVLSYK LNAVAEAVLK  
25 551 DKKKDLRYD IPAYYASGPA QRGVIGEYCV QDSSLVGQLF FKFLPHLELS  
601 AVARLAGINI TRTIYDQQI RVFTCLRLA GQKGFLPDT QGRFRGLDKE  
651 APKRAVPRG EGERPGDGNG DEDKDDDEDE DGDEREEVAR ETGGRHVGYQ  
30 701 GARVLDPTSG FHVDPVVVFD FASLYPSIIQ AHNLCFSTLS LRPEAVAHLE  
751 ADRDYLEIEV GGRRLFFVKA HVRESLLSIL LRDWLAMRKQ IRSRIPQSTP  
35 801 EEAVLLDKQQ AAIKVCNSV YGFTGAQHGL LPCLHVAATV TTIGREMLLA  
851 TRAYVHARWA EFDQLLADFP EAAGMRAPGP YSMRIYGDT DSIFVLCRGL  
901 TAAGLVAMGD KMASHISRAL FLPIKLECE KTFTKLLIA KKKYIGVICG  
40 951 GKMLIKGVDL VRKNNCAFIN RTSRALVDLL FYDDTVSGAA AALAERPAEE  
1001 WLARPLPEGL QAFGAVLVDA HRRITDPERD IQDFVLTAE SRHPRAYTNK  
45 1051 RLAHLTVYYK LMARRAQVPS IKDRIPYVIV AQTREVEETV ARLAALRELD  
1101 AAAPGDEPAP PAALPSPAKR PRETPSHADP PGGASKPRKL LVSELAEDPG  
1151 YAIARGVPLN TDYYFSHLLG AACVTFKALF GNNAKITESL LKRFPETWH  
50 1201 PPDDVAARLR AAGFGPAGAG ATAEETRRML HRAFDTLA\*

**SEQ.ID.NO. 3** DNA sequence of DNA polymerase gene for HSV2-186-M1

1 ATGTTTTGTG CCGCGGGCGG CCCGGCTTCC CCCGGGGGGA AGTCGGCGGC  
5 51 TCGGGCGGCG TCTGGGTTTT TTGCCCCCA CAACCCCGG GGAGCCACCC  
101 AGACGGCACC GCCGCCTTGC CGCCGGCAGA ACTTCTACAA CCCCCACCTC  
151 GCTCAGACCG GAACGCAGCC AAAGGCCCCC GGGCCGGCTC AGCGCCATAC  
10 201 GTACTACAGC GAGTGCAGC AATTTCGATT TATCGCCCCG CGTTCGCTGG  
251 ACGAGGACGC CCCCGCGGAG CAGCGCACCG GGGTCCACGA CGGCCGCCTC  
15 301 CGGCGCGCCC CTAAGGTGTA CTGCGGGGGG GACGAGCGCG ACGTCTCCG  
351 CGTGGGCCCC GAGGGCTTCT GGCCGCGTCG CTTGCGCCTG TGGGGCGGTG  
401 CGGACCATGC CCCCAGGGG TTCGACCCA CCGTCACCGT CTTCCACGTG  
20 451 TACGACATCC TGGAGCACGT GGAACACGCG TACAGCATGC GCGCCGCCCA  
501 GCTCCACGAG CGATTATGG ACGCCATCAC GCCCGCCGGG ACCGTCATCA  
25 551 CGCTTCTGGG TCTGACCCCC GAAGGCCATC GCGTCGCCGT TCACGTCTAC  
601 GGCACGCGGC AGTACTTTTA CATGAACAAG GCGGAGGTGG ATCGGCACCT  
651 GCAGTGCCGT GCCCCGCGCG ATCTCTGCGA GCGCCTGGCG GCGGCCCTGC  
30 701 GCGAGTCGCC GGGGGCGTCG TTCCGCGGCA TCTCCGCGGA CCACTTCGAG  
751 GCGGAGGTGG TGGAGCGCGC CGACGTGTAC TATTACGAAA CGCGCCCGAC  
35 801 CCTGTACTAC CGCGTCTTCG TCGAAGCGG GCGCGCGCTG GCCTACCTGT  
851 GCGACAACTT TTGCCCCGCG ATCAGGAAGT ACGAGGGGGG CGTCGACGCC  
901 ACCACCCGGT TTATCCTGGA CAACCCGGGG TTTGTACCT TCGGCTGGTA  
40 951 CCGCCTCAAG CCCGGCCGCG GGAACGCGCC GGCCCAACCG CGCCCCCGA  
1001 CGGCGTTCGG AACCTCGAGC GACGTCGAGT TTAAGTGCAC GGCGGACAAC  
45 1051 CTGGCCGTCG AGGGGGCCAT GTGTGACCTG CCGGCCTACA AGCTCATGTG  
1101 CTTGATATC GAATGCAAGG CCGGGGGGGA GGACGAGCTG GCCTTTCGG  
1151 TCGCGGAACG CCCGGAAGAC CTCGTCATCC AGATCTCCTG TCTGCTCTAC  
50 1201 GACCTGTCCA CCACCGCCCT CGAGCACATC CTCCTGTTTT CGCTCGGATC  
1251 CTGCGACCTC CCCGAGTCCC ACCTCAGCGA TCTCGCCTCC AGGGGCCTGC  
55 1301 CGGCCCCCGT CGTCTTGGAG TTTGACAGCG AATTCGAGAT GCTGCTGGCC  
1351 TTCATGACCT TCGTCAAGCA GTACGGCCCC GAGTTCGTGA CCGGGTACAA  
1401 CATCATCAAC TTCGACTGGC CCTTCGTCCT GACCAAGCTG ACGGAGATCT  
60

1451 ACAAGGTCCC GCTCGACGGG TACGGGCGCA TGAACGGCCG GGGTGTGTTC  
1501 CGCGTGTGGG ACATCGGCCA GAGCCACTTT CAGAAGCGCA GCAAGATCAA  
5 1551 GGTGAACGGG ATGGTGAACA TCGACATGTA CGGCATCATC ACCGACAAGG  
1601 TCAAACCTCTC CAGCTACAAG CTGAACGCCG TCGCCGAGGC CGTCTTGAAG  
1651 GACAAGAAGA AGGATCTGAG CTACCGCGAC ATCCCCGCCT ACTACGCCTC  
10 1701 CGGGCCCCGCG CAGCGCGGGG TGATCGGCGA GTATTGTGTG CAGGACTCGC  
1751 TGCTGGTCGG GCAGCTGTTC TTCAAGTTTC TGCCGCACCT GGAGCTTTCC  
1801 GCCGTCGCGC GCCTGGCGGG CATCAACATC ACCCGCACCA TCTACGACGG  
1851 CCAGCAGATC CGCGTCTTCA CGTGCTCCT GCGCCTTGCG GGCCAGAAGG  
1901 GCTTCATCCT GCCGGACACC CAGGGGCGGT TTCGGGGCCT CGACAAGGAG  
20 1951 GCGCCCAAGC GCCCGGCCGT GCCTCGGGGG GAAGGGGAGC GGCCGGGGGA  
2001 CGGGAACGGG GACGAGGATA AGGACGACGA CGAGGACGGG GACGAGGACG  
25 2051 GGGACGAGCG CGAGGAGGTC GCGCGCGAGA CCGGGGGCCG GCACGTTGGG  
2101 TACCAGGGGG CCCGGGTCCT CGACCCACC TCCGGGTTTC ACGTCGACCC  
2151 CGTGGTGGTG TTTGACTTTG CCAGCCTGTA CCCCAGCATC ATCCAGGCCC  
30 2201 ACAACCTGTG CTTAGTACG CTCTCCCTGC GGCCCGAGGC CGTCGCGCAC  
2251 CTGGAGGCGG ACCGGGACTA CCTGGAGATC GAGGTGGGGG GCCGACGGCT  
35 2301 GTTCTTCGTG AAGGCCACG TACGCGAGAG CCTGCTGAGC ATCCTGCTGC  
2351 GCGACTGGCT GGCCATGCGA AAGCAGATCC GCTCGCGGAT CCCCCAGAGC  
2401 CCCCCGAGG AGGCCGTCCT CCTCGACAAG CAACAGGCCG CCATCAAGGT  
40 2451 GGTGTGCAAC TCGGTGTACG GGTTACCGG GGCGCAGCAC GGTCTTCTGC  
2501 CCTGCCTGCA CGTGGCCGCC ACCGTGACGA CCATCGGCCG CGAGATGCTC  
45 2551 CTCGCGACGC GCGCGTACGT GCACGCGCGC TGGGCGGAGT TCGATCAGCT  
2601 GCTGGCCGAC TTTCCGGAGG CGGCCGGCAT GCGCGCCCCC GGTCCGTACT  
2651 CCATGCGCAT CATCTACGGG GACACGGA CTATTTTCGT TTTGTGCCGC  
50 2701 GGCCTACGG CCGCGGGCCT GGTGGCCATG GGCGACAAGA TGGCGAGCCA  
2751 CATCTCGCGC GCGCTGTTCC TCCCCCGAT CAAGCTCGAG TGCGAAAAA  
55 2801 CGTTCACCAA GCTGCTGCTC ATCGCCAAGA AAAAGTACAT CGGCGTCATC  
2851 TGCGGGGGCA AGATGCTCAT CAAGGGCGTG GATCTGGTGC GCAAAAACAA  
2901 CTGCGCGTTT ATCAACCGCA CCTCCAGGGC CCTGGTCGAC CTGCTGTTTT  
60

2951 ACGACGATAC CGTATCCGGA GCGGCCGCCG CGTTAGCCGA GCGCCCCGCA  
3001 GAGGAGTGGC TGGCGCGACC CCTGCCCCGAG GGA CTGCAGG CGTTCGGGGC  
5 3051 CGTCCTCGTA GACGCCCATC GGCGCATCAC CGACCCGGAG AGGGACATCC  
3101 AGGACTTTGT CCTCACCGCC GAACTGAGCA GACACCCGCG CGCGTACACC  
3151 AACAAGCGCC TGGCCCACCT GACGGTGTAT TACAAGCTCA TGGCCCCGCC  
10 3201 CGCGCAGGTC CCGTCCATCA AGGACCGGAT CCCGTACGTG ATCGTGGCCC  
3251 AGACCCGCGA GG TAGAGGAG ACGGTCGCGC GGCTGGCCGC CCTCCGCGAG  
15 3301 CTAGACGCCG CCGCCCCAGG GGACGAGCCC GCCCCCCCAG CGGCCCTGCC  
3351 CTCCCCGCC AAGCGCCCC GGGAGACGCC GTCGCATGCC GACCCCCCGG  
3401 GAGGCGCGTC CAAGCCCCGC AAGCTGCTGG TGTCCGAGCT GGCGGAGGAT  
20 3451 CCCGGGTACG CCATCGCCCC GGGCGTTCCG CTCAACACGG ACTATTACTT  
3501 CTCGCACCTG CTGGGGGCGG CCTGCGTGAC GTTCAAGGCC CTGTTTGAA  
25 3551 ATAACGCCAA GATCACCGAG AGTCTGTTAA AGAGGTTTAT TCCCGAGACG  
3601 TGGCACCCCC CGGACGACGT GGCCGCGCGG CTCAGGGCCG CGGGGTTCGG  
3651 GCCGGCGGGG GCCGGCGCTA CGGCGGAGGA AACTCGTCGA ATGTTGCATA  
30 3701 GAGCCTTTGA TACTCTAGCA TGA

## SEQ.ID.NO. 4

Amino acid sequence of DNA polymerase for HSV2-186-M1

5       1 MFCAAGGPAS PGGKSAARAA SGFFAPHNPR GATQTAPPPC RRQNFYNPHL  
      51 AQTGTQPKAP GPAQRHTYYS ECDEFRIAP RSLDEDAPAE QRTGVHDGRL  
      101 RRAPKVYCGG DERDVLRVGP EGFWRRLRL WGGADHAPEG FDPTVTVFHV  
10       151 YDILEHVEHA YSMRAAQLHE RFMDAITPAG TVITLLGLTP EGHRAVAVHVY  
      201 GTRQFYFMNK AEVDRHLQCR APRDLCERLA AALRESPGAS FRGISADHFE  
      251 AEVVERADVY YYETRPTLYY RVFVRSGRAL AYLCDNFCPA IRKYEGGVDA  
15       301 TTRFILDNPG FVTFGWYRLK PGRGNAPAQP RPPTAFGTSS DVEFNCTADN  
      351 LAVEGAMCDL PAYKLMCFDI ECKAGGEDEL AFPVAERPED LVIQISCLLY  
20       401 DLSTTALEHI LLFSLGSCDL PESHLSDLAS RGLPAPVVLE FDSEFEMLLA  
      451 FMTFVKQYGP EFVTGYNIIN FDWPFVLTCL TEIYKVPLDG YGRMNNGRVF  
      501 RVWDIGQSHF QKRSEIKVNG MVNIDMYGII TDKVKLSSYK LNAVAEAVLK  
25       551 DKKKDLSTYR IPAYYASGPA QRGVIGEYCV QDSSLVGQLF FKFLPHLELS  
      601 AVARLAGINI TRTTYDGQOI RVFTCLRLA GQKGFILPDT QGRFRGLDKE  
30       651 APKRAVPRG EGERPGDGNG DEDKDDDEDG DEDGDEREEV ARETGGRHVG  
      701 YQGARVLDPT SGFHVDPVVV FDFASLYPSI IQAHNLCFST LSLRPEAVAH  
      751 LEADRDYLEI EVGGRRLLFFV KAHVRESLLS ILLRDWLAMR KQIRSRIPQS  
35       801 PPEEAVLLDK QQAAIKVVCN SVYGFTGAQH GLLPCLHVAA TVTTIGREML  
      851 LATRAYVHAR WAEFDQLLAD FPEAAGMRAP GPYSMRITYG DTDSIFVLCR  
40       901 GLTAAGLVAM GDKMASHISR ALFLPPIKLE CEKTFTKLLL IAKKKYIGVI  
      951 CGGKMLIKGV DLVRKNNCAF INRTSRALVD LLFYDDTVSG AAAALAERPA  
45       1001 EEWLARPLPE GLQAFGAVLV DAHRRITDPE RDIQDFVLTA ELSRHPRAYT  
      1051 NKRLAHLTVY YKLMARRAQV PSIKDRIPYV IVAQTREVEE TVARLAALRE  
      1101 LDAAAPGDEP APPAALPSA KRPRETPSHA DPPGGASKPR KLLVSELAED  
50       1151 PGYAIARGVP LNTDYYFSLH LGAACVTFKA LFGNNAKITE SLLKRFIPET  
      1201 WHPPDDVAAR LRAAGFGPAG AGATAEETRR MLHRAFDTLA \*

## SEQ.ID.NO. 5 DNA sequence of DNA polymerase gene for HSV1-KOS-M1

1 ATGTTTTCCG GTGGCGGCGG CCCGCTGTCC CCCGGAGGAA AGTCGGCGGC  
5 51 CAGGGCGGCG TCCGGGTTTT TTGCGCCCGC CGGCCCTCGC GGAGCCGGCC  
101 GGGGACCCCC GCCTTGTTTG AGGCAAACT TTTACAACCC CTACCTCGCC  
151 CCAGTCGGGA CGCAACAGAA GCCGACCGGG CCAACCCAGC GCCATACGTA  
10 201 CTATAGCGAA TGCATGAAT TTCGATTCAT CGCCCCGCGG GTGCTGGACG  
251 AGGATGCCCC CCCGGAGAAG CGCGCCGGGG TGCACGACGG TCACCTCAAG  
15 301 CGCGCCCCA AGGTGTACTG CGGGGGGGAC GAGCGCGACG TCCTCCGCGT  
351 CGGGTCGGGC GGCTTCTGGC CGCGGCGCTC GCGCCTGTGG GGCGGCGTGG  
401 ACCACGCCCC GCGGGGGTTC AACCCACCG TCACCGTCTT TCACGTGTAC  
20 451 GACATCCTGG AGAACGTGGA GCACGCGTAC GGCATGCGCG CGGCCAGTT  
501 CCACGCGCGG TTTATGGACG CCATCACACC GACGGGGACC GTCATCACGC  
25 551 TCCTGGGCCT GACTCCGGAA GGCCACCGGG TGGCCGTTC ACGTTACGGC  
601 ACGCGGCAGT ACTTTTACAT GAACAAGGAG GAGGTTGACA GGCACCTACA  
651 ATGCCGCGCC CCACGAGATC TCTGCGAGCG CATGGCCGCG GCCCTGCGCG  
30 701 AGTCCCCGGG CGCGTCGTT CGCGGCATCT CCGCGGACCA CTTCGAGGCG  
751 GAGGTGGTGG AGCGCACCGA CGTGTACTAC TACGAGACGC GCCCCGCTCT  
35 801 GTTTTACCGC GTCTACGTCC GAAGCGGGCG CGTGCTGTCG TACCTGTGCG  
851 ACAACTTCTG CCCGGCCATC AAGAAGTACG AGGGTGGGGT CGACGCCACC  
901 ACCCGGTTCA TCCTGGACAA CCCCGGGTTC GTCACCTTCG GCTGGTACCG  
40 951 TCTCAAACCG GGCCGGAACA ACACGCTAGC CCAGCCGCGG GCCCCGATGG  
1001 CCTTCGGGAC ATCCAGCGAC GTCGAGTTTA ACTGTACGGC GGACAACCTG  
45 1051 GCCATCGAGG GGGGCATGAG CGACCTACCG GCATACAAGC TCATGTGCTT  
1101 CGATATCGAA TGCAAGGCGG GGGGGGAGGA CGAGCTGGCC TTTCCGGTGG  
1151 CCGGGCACCC GGAGGACCTG GTTATTCAGA TATCCTGTCT GCTCTACGAC  
50 1201 CTGTCCACCA CCGCCCTGGA GCACGTCCTC CTGTTTTCGC TCGGTTCCCTG  
1251 CGACCTCCCC GAATCCACCG TGAACGAGCT GGCGGCCAGG GGCCTGCCCA  
55 1301 CGCCCCGTGGT TCTGGAATTC GACAGCGAAT TCGAGATGCT GTTGGCCTTC  
1351 ATGACCCTTG TGAAACAGTA CGGCCCCGAG TTCGTGACCG GGTACAACAT  
1401 CATCAACTTC GACTGGCCCT TCTTGCTGGC CAAGTTGACG GACATTTACA  
60

1451 AGGTCCCCCT GGACGGGTAC GGCCGCATGA ACGGCCGGGG CGTGTTTCGC  
1501 GTGTGGGACA TAGGCCAGAG CCACTTCCAG AAGCGCAGCA AGATAAAGGT  
5 1551 GAACGGCATG GTGAACATCG ACATGTACGG GATCATAACC GACAAGATCA  
1601 AGCTCTCGAG CTACAAGCTC AACGCCGTGG CCGAAGCCGT CCTGAAGGAC  
1651 AAGAAGAAGG ACCTGAGCTA TCGCGACATC CCCGCCTACT ACGCCGCCGG  
10 1701 GCCCGCGCAA CGCGGGGTGA TCGCGAGTA CTGCATACAG GATTCCCTGC  
1751 TGGTGGGCCA GCTGTTTTTT AAGTTTTTGC CCCATCTGGA GCTCTCGGCC  
1801 GTCGCGCGCT TGGCGGGTAT TAACATCACC CGCACCATCT ACGACGGCCA  
15 1851 GCAGATCCGC GTCTTTACGT GCCTGCTGCG CCTGGCCGAC CAGAAGGGCT  
1901 TTATTCTGCC GGACACCCAG GGGCGATTTA GGGGCGCCGG GGGGAGGCG  
20 1951 CCCAAGCGTC CGGCCGAGC CCGGGAGGAC GAGGAGCGGC CAGAGGAGGA  
2001 GGGGAGGAC GAGGACGAAC GCGAGGAGGG CGGGGGCGAG CGGGAGCCGG  
2051 AGGGCGCGCG GGAGACCGCC GGCCGGCACG TGGGGTACCA GGGGGCCAGG  
25 2101 GTCCTTGACC CCACTTCCGG GTTTCACGTG AACCCCGTGG TGGTGTTCTGA  
2151 CTTTGCCAGC CTGTACCCCA GCATCATCCA GGCCCACAAC CTGTGCTTCA  
30 2201 GCACGCTCTC CCTGAGGGCC GACGCAGTGG CGCACCTGGA GGCGGGCAAG  
2251 GACTACCTGG AGATCGAGGT GGGGGGGCGA CGGCTGTTCT TCGTCAAGGC  
35 2301 TCACGTGCGA GAGAGCCTCC TCAGCATCCT CCTGCGGGAC TGGCTCGCCA  
2351 TGCGAAAGCA GATCCGCTCG CGGATTCCCC AGAGCAGCCC CGAGGAGGCC  
2401 GTGCTCCTGG ACAAGCAGCA GGCCGCCATC AAGGTCGTGT GTAACCTGGT  
40 2451 GTACGGGTTC ACGGGAGCGC AGCACGGA CTGCGCGTGC CTGCACGTTG  
2501 CCGCGACGGT GACGACCATC GGCCGCGAGA TGCTGCTCGC GACCCGCGAG  
45 2551 TACGTCCACG CGCGCTGGGC GGCCITCGAA CAGCTCCTGG CCGATTTCCT  
2601 GGAGGCGGCC GACATGCGCG CCCCCGGGCC CTATTCCATG CGCATCATCT  
2651 ACGGGGACAC GGA CTCCATA TTTGTGCTGT GCCGCGGCCT CACGGCCGCC  
50 2701 GGGCTGACGG CCATGGGCGA CAAGATGGCG AGCCACATCT CGCGCGCGCT  
2751 GTTCTGCCC CCCATCAAAC TCGAGTGCGA AAAGACGTTT ACCAAGCTGC  
2801 TGCTGATCGC CAAGAAAAAG TACATCGGCG TCATCTACGG GGGTAAGATG  
2851 CTCATCAAGG GCGTGGATCT GGTGCGCAAA AACAACTGCG CGTTTATCAA  
2901 CCGCACCTCC AGGGCCCTGG TCGACCTGCT GTTTTACGAC GATACCGTAT  
60

2951 CCGGAGCGGC CGCCGCGTTA GCCGAGCGCC CCGCAGAGGA GTGGCTGGCG  
3001 CGACCCCTGC CCGAGGGACT GCAGGCGTTC GGGGCCGTCC TCGTAGACGC  
5 3051 CCATCGGCGC ATCACCGACC CGGAGAGGGA CATCCAGGAC TTTGTCCTCA  
3101 CCGCCGA ACT GAGCAGACAC CCGCGCGCGT ACACCAACAA GCGCCTGGCC  
3151 CACCTGACGG TGTATTACAA GTCATGGCC CGCCGCGCGC AGGTCCCGTC  
10 3201 CATCAAGGAC CGGATCCCGT ACGTGATCGT GGCCAGACC CGCGAGGTAG  
3251 AGGAGACGGT CGCGCGGCTG GCCGCCCTCC GCGAGCTAGA CGCCGCCGCC  
15 3301 CCAGGGGACG AGCCCGCCCC CCGCGCGGCC CTGCCCTCCC CGGCCAAGCG  
3351 CCCCCGGGAG ACGCCGTCGC ATGCCGACCC CCGGGAGGC GCGTCCAAGC  
3401 CCCGCAAGCT GCTGGTGTCC GAGCTGGCCG AGGATCCCGC ATACGCCATT  
20 3451 GCCCACGGCG TCGCCCTGAA CACGGA CTAT TACTTCTCCC ACCTGTTGGG  
3501 GCGGGCGTGC GTGACATTCA AGGCCCTGTT TGGGAATAAC GCCAAGATCA  
25 3551 CCGAGAGTCT GTTAAAAAGG TTTATTCCCG AAGTGTGGCA CCCCCGGAC  
3601 GACGTGGCCG CGCGGCTCCG GGCCGAGGG TTCGGGGCGG TGGGTGCCGG  
3651 CGCTACGGCG GAGGAACTC GTCGAATGTT GCATAGAGCC TTTGATACTC  
30 3701 TAGCATGA

**SEQ.ID.NO. 6** Amino acid sequence of DNA polymerase for HSV1-KOS-M1

1 MFSGGGGPLS PGGKSAARAA SGFFAPAGPR GAGRGGPPCL RQNFYNPYLA  
5 51 PVGTQKPTG PTQRHTYYSE CDEFRFIAPR VLDEDAPEK RAGVHDGHLK  
101 RAPKVYCGGD ERDVLRVGSG GFWPRRSRLW GGVDHAPAGF NPTVTVFHVY  
10 151 DILENVEHAY GMRAAQFHAR FMDAITPTGT VITLLGLTPE GHRVAHVHYG  
201 TRQYFYMNKE EVDRHLQCRA PRDLCERMAA ALRESPGASF RGISADHFEA  
251 EVVERTDVYY YETRPALFYR VYVRSGRVLS YLCDNFCPAI KKYEGGV DAT  
15 301 TRFILDNPGF VTFGWYRLKP GRNNTLAQPR APMAFGTSSD VEFNCTADNL  
351 AIEGGMSDLP AYKLMCFDIE CKAGGEDELA FPVAGHPEDL VIQISCLLYD  
20 401 LSTTALEHVL LFSLGSCDLP ESHLNELAAR GLPTPVVLEF DSEFEMLLAF  
451 MTLVKQYGPE FVTGYNINIF DWPFLAKLT DIYKVPLDGY GRMNGRGVFR  
501 VWDIGQSHFQ KRSKIKVNGM VNIDMYGIIT DKIKLSSYKL NAVA EAVLKD  
25 551 KKKDLSYRDI PAYYAAGPAQ RGVIGEYCIQ DSVLVGQLFF KFLPHLELSA  
601 VARLAGINT RTTYDGQIR VFTCLRLAD QKGFILPDTQ GRFRGAGGEA  
30 651 PKRPAAARED EERPEEEGED EDEREEGGGE REPEGARETA GRHVG YQGAR  
701 VLDPTSGFHV NPVVVFDFAS LYPSIIQAHN LCFSTLSLRA DAVA HLEAGK  
751 DYLEIEVGR RLFFVKAHVR ESLLSILLRD WLAMRKQIRS RIPQSSPEEA  
35 801 VLLDKQAAI KVCNSVYGF TGAQHGLLPC LHVAATVTTI GREMLLATRE  
851 YVHARWAAFE QLLADFPEAA DMRAPGPYSM RIYGD TDSI FVLCRGLTAA  
40 901 GLTAMGDKMA SHISRALFLP PIKLECEKTF TKLLLIAKKK YIGVIYGGKM  
951 LIKGVDLVRK NNCAFINRTS RALVDLLFYD DTVSGAAAAL AERPAEEWLA  
1001 RLPPEGLQAF GAVLVDAHRR ITDPERDIQD FVLTAELSRH PRAYTNKRLA  
45 1051 HLTVYYKLMA RRAQVPSIKD RPYVIVAQT REVEETVARL AALRELDAAA  
1101 PGDEPAPPAA LPSPAKRPRE TP SHADPPGG ASKPRKLLVS ELAEDPAYAI  
50 1151 AHGVALNTDY YFSHLLGAAC VTFKALFGNN AKITESLLKR FIPEVWHPPD  
1201 DVAARLRAAG FGAVGAGATA EETRRMLHRA FDTLA\*

## SEQ.ID.NO. 7 DNA sequence of HSV polymerase gene for HSV1-F-M1

```

5      1  ATGTTTTCCTG GTGGCGGCGG CCCGCTGTCC CCCGGAGGAA AGTCGGCGGC
      51  CAGGGCGGCG TCCGGGTTTT TTGCGCCCGC CGGCCCTCGC GGAGCCGGCC
     101  GGGGACCCCC GCCTTGCTTG AGGCAAAACT TTTACAACCC CTACCTCGCC
     151  CCAGTCGGGA CGCAACAGAA GCCGACCGGG CCAACCCAGC GCCATACGTA
     201  CTATAGCGAA TGCGATGAAT TTCGATTCAT CGCCCCGCGG GTGCTGGACG
     251  AGGATGCCCC CCCGGAGAAG CGCGCCGGGG TGACGACGCG TCACCTCAAG
     301  CGCGCCCCCA AGGTGTACTG CGGGGGGGAC GAGCGCGACG TCCTCCGCGT
     351  CGGGTCGGGC GGCTTCTGGC CGCGGCGCTC GCGCCTGTGG GGC GCGTGG
     401  ACCACGCCCC GCGGGGGTTC AACCCACCG TCACCGTCTT TCACGTGTAC
     451  GACATCCTGG AGAACGTGGA GCACGCGTAC GGCATGCGCG CGGCCAGTT
     501  CCACGCGCGG TTTATGGACG CCATCACACC GACGGGGACC GTCATCACGC
     551  TCCTGGGCCT GACTCCGGA GGCCACCGGG TGGCCGTTCA CGTTTACGGC
     601  ACGCGGCAGT ACTTTTACAT GAACAAGGAG GAGGTCGACA GGCACCTACA
     651  ATGCCGCGCC CCACGAGATC TCTGCGAGCG CATGGCCGCG GCCCTGCGCG
     701  AGTCCCCGGG CGCGTCGTTC CGCGGCATTT CCGCGGACCA CTTCGAGGCG
     751  GAGGTGGTGG AGCGCACCGA CGTGACTACT TACGAGACGC GCCCGCTCT
     801  GTTTTACCGC GTCTACGTCC GAAGCGGGCG CGTGCTGTCT TACCTGTGCG
     851  ACAACTTCTG CCCGGCCATC AAGAAGTACG AGGGTGGGGT CGACGCCACC
     901  ACCCGGTTCA TCCTGGACAA CCCC GGTTTC GTACCTTCG GCTGGTACCG
     951  TCTCAAACCG GGCCGGAACA ACACGCTAGC CCAGCCGCGG GCGCCGATGG
    1001  CCTTCGGGAC ATCCAGCGAC GTCGAGTTTA ACTGTACGGC GGACAACCTG
    1051  GCCATCGAGG GGGGCATGAG CGACCTACCG GCATACAAGC TCATGTGCTT
    1101  CGATATCGAA TGCAAGGCGG GGGGGGAGGA CGAGCTGGCC TTTCCGGTGG
    1151  CCGGGCACCC GGAGGACCTG GTCATCCAGA TATCCTGTCT GCTCTACGAC
    1201  CTGTCCACCA CCGCCCTGGA GCACGTCCTC CTGTTTTCGC TCGGTTCCCTG
    1251  CGACCTCCCC GAATCCCACC TGAACGAGCT GGCGGCCAGG GGCCTGCCCA
    1301  CGCCCGTGGT TCTGGAATTC GACAGCGAAT TCGAGATGCT GTTGGCCTTC
    1351  ATGACCCTTG TGAAACAGTA CGGCCCCGAG TTCGTGACCG GGTACAACAT
    1401  CATCAACTTC GACTGGCCCT TCTTGCTGGC CAAGCTGACG GACATTTACA
    1451  AGGTCCCCCT GGACGGGTAC GGCCGCATGA ACGGCCGGGG CGTGTTTCGC
    1501  GTGTGGGACA TAGGCCAGAG CCACTTCCAG AAGCGCAGCA AGATAAAGGT
    1551  GAACGGCATG GTGAACATCG ACATGTACGG GATTATAACC GACAAGATCA

```

1601 AGCTCTCGAG CTACAAGCTC AACGCCGTGG CCGAAGCCGT CCTGAAGGAC  
 5 1651 AAGAAGAAGG ACCTGAGCTA TCGCGACATC CCCGCCTACT ACGCCGCCGG  
 1701 GCCCGCGCAA CGCGGGGTGA TCGCGAGTA CTGCATACAG GATTCCCTGC  
 1751 TGGTGGGCCA GCTGTTTTTT AAGTTTTTGC CCCATCTGGA GCTCTCGGCC  
 10 1801 GTCGCGCGCT TGGCGGGTAT TAACATCACC CGCACCATCT ACGACGGCCA  
 1851 GCAGATCCGC GTCTTTTACGT GCCTGCTGCG CCTGGCCGAC CAGAAGGGCT  
 15 1901 TTATTCTGCC GGACACCCAG GGGCGATTTA GGGGCGGCGG GGGGGAGGCG  
 1951 CCCAAGCGTC CGGCCGCAGC CCGGGAGGAC GAGGAGCGGC CAGAGGAGGA  
 2001 GGGGGAGGAC GAGGACGAAC GCGAGGAGGG CGGGGGCGAG CGGGAGCCGG  
 20 2051 AGGGCGCGCG GGAGACCGCC GGCCGGCACG TGGGGTACCA GGGGGCCAGG  
 2101 GTCCTTGACC CCACTTCCGG GTTTCATGTG AACCCCGTGG TGGTGTTCGA  
 2151 CTTTGCCAGC CTGTACCCCA GCATCATCCA GGCCACAAAC CTGTGCTTCA  
 25 2201 GCACGCTCTC CCTGAGGGCC GACGCAGTGG CGCACCCTGA GGCGGGCAAG  
 2251 GACTACCTGG AGATCGAGGT GGGGGGGCGA CGGCTGTTCT TCGTCAAGGC  
 30 2301 TCACGTGCGA GAGAGCCTCC TCAGCATCCT CCTGCGGGAC TGGCTCGCCA  
 2351 TCGCAAAGCA GATCCGCTCG CGGATTCCCC AGAGCAGCCC CGAGGAGGCC  
 2401 GTGCTCCTGG ACAAGCAGCA GGCCGCCATC AAGGTCGTGT GTAACTCGGT  
 35 2451 TTACGGGTTC ACGGGAGCGC AGCACGGACT CCTGCCGTGC CTGCACGTTG  
 2501 CCGCGACGGT GACGACCATC GGCCGCGAGA TGCTGCTCGC GACCCGCGAG  
 40 2551 TACGTCCACG CGCGCTGGGC GGCCTTCGAA CAGCTCCTGG CCGATTTCCT  
 2601 GGAGGCGGCC GACATGCGCG CCCCCGGGCC CTATTCCATG CGCATCATCT  
 2651 ACGGGGACAC GGAATCCATC TTTGTGCTGT GCCGCGGCCT CACGGCCGCC  
 45 2701 GGGCTGACGG CCGTGGGCGA CAAGATGGCG AGCCACATCT CGCGCGCGCT  
 2751 GTTTCTGTCC CCCATCAAAC TCGAGTGC GAAGACGTTT ACCAAGCTGC  
 50 2801 TGCTGATCGC CAAGAAAAAG TACATCGGCG TCATCTACGG GGGTAAGATG  
 2851 CTCATCAAGG GCGTGGATCT GGTGCGCAA AACAACTGCG CGTTTATCAA  
 2901 CCGCACCTCC AGGGCCCTGG TCGACCTGCT GTTTTACGAC GATACCGTAT  
 55 2951 CCGGAGCGGC CGCCGCGTTA GCCGAGCGCC CCGCAGAGGA GTGGCTGGCG  
 3001 CGACCCCTGC CCGAGGGACT GCAGGCGTTC GGGGCCGTCC TCGTAGACGC  
 60 3051 CCATCGGCGC ATCACCGACC CGGAGAGGGA CATCCAGGAC TTTGTCTCTA  
 3101 CCGCCGAAC T GAGCAGACAC CCGCGCGCGT ACACCAACAA GCGCCTGGCC  
 3151 CACCTGACGG TGTATTACAA GCTCATGGCC CGCCGCGCGC AGGTCCCCTC  
 65 3201 CATCAAGGAC CGGATCCCGT ACGTGATCGT GGCCAGACC CGCGAGGTAG

3251 AGGAGACGGT CGCGCGGCTG GCCGCCCTCC GCGAGCTCGA CGCCGCCGCC  
3301 CCAGGGGACG AGCCCGCCCC CCCC GCGGCC CTGCCCTCCC CGGCCAAGCG  
5 3351 CCCCCGGGAG ACGCCGTTGC ATGCCGACCC CCCGGGAGGC GCGTCCAAGC  
3401 CCCGCAAGCT GCTGGTGTCC GAGCTGGCCG AGGATCCCGC ATACGCCATT  
3451 GCCCACGGCG TCGCCCTGAA CACGGACTAT TACTTCTCCC ACCTGTTGGG  
10 3501 GCGGGCGTGC GTGACATTCA AGGCCCTGTT TGGGAATAAC GCCAAGATCA  
3551 CCGAGAGTCT GTTAAAAAGG TTTATTCCCG AAGTGTGGCA CCCCCGGAC  
15 3601 GACGTGGCCG CGCGGCTCCG GGCCGCAGGG TTCGGGGCGG TGGGTGCCCG  
3651 CGCTACGGCG GAGGAAACTC GTCGAATGTT GCATAGAGCC TTTGATACTC  
3701 TAGCATGA

## SEQ.ID.NO. 8 Amino acid sequence of DNA polymerase for HSV1-F-M1

1 MFSGGGGPLS PGGKSAARAA SGFFAPAGPR GAGRGPPPCL RQNFYNPYLA  
5 51 PVGTQQKPTG PTQRHTYYSE CDEFRFIAPR VLDEDAPPEK RAGVHDGHLK  
101 RAPKVYCGGD ERDVLRVGSG GFWPRRSRLW GGVDHAPAGF NPTVTVFHVY  
151 DILENVEHAY GMRAAQFHAR FMDAITPTGT VITLLGLTPE GHRVAVHVYG  
10 201 TRQYFYMNKE EVDRHLQCRA PRDLCERMAA ALRESPGASF RGISADHFEA  
251 EVVERTDVYY YETRPALFYR VYVRSGRVLS YLCDNFCPAI KKYEGGV DAT  
15 301 TRFILDNPGF VTFGWYRLKP GRNNTLAQPR APMAFGTSSD VEFNCTADNL  
351 AIEGGMSDLP AYKLMCFDIE CKAGGEDELA FPVAGHPEDL VIQISCLLYD  
401 LSTTALEHVL LFSLGSCDLP ESHLNELAAR GLPTPVVLEF DSEFEMLLAF  
20 451 MTLVKQYGPE FVTGYNIINF DWPFLAKLT DIYKVPLDGY GRMN GRGVFR  
501 VWDIGQSHFQ KRSKIKVNGM VNIDMYGIIT DKIKLSSYKL NAVAEAVLKD  
25 551 KKKDLSYRDI PAYYAAGPAQ RGVIGEYCIQ DSVLVGQLFF KFLPHLELSA  
601 VARLAGINIT RTTYDGQQIR VFTCLRLAD QKGFIPLDTQ GRFRGGGGEEA  
651 PKRPAAAREE EERPEEEGED EDEREEGGGE REPEGARETA GRHVG YQGAR  
30 701 VLDPTSGFHV NPVVVDFAS LYPSTQAHN LCFSTLSLRA DAVAHLEAGK  
751 DYLEIEVGGR RLFFVKAHVR ESLLSILLRD WLAMRKQIRS RIPQSSPEEA  
35 801 VLLDKQQA AI KVV CNSVYGF TGAQHGLLPC LHVAATVT TI GREMLLATRE  
851 YVHARWAAFE QLLADFPEAA DMRAPGPYSM RIYGD TDSI FVLCRGLTAA  
901 GLTAVGDKMA SHISRALFLS PIKLECEKTF TKLLLI AKKK YIGVIYGGKM  
40 951 LIKGVDLVRK NNCAFINRTS RALVDLLFYD DTVSGAAAAL AERP AEWLA  
1001 RPLPEGLQAF GAVLVDAHRR ITDPERDIQD FVLTAELSRH PRAYTNKRLA  
45 1051 HLTVYYKLMA RRAQVPSIKD RPYVIVAQT REVEETVARL AALRELDAAA  
1101 PGDEPAPPAA LPSPAKRPRE TPLHADPPGG ASKPRKLLVS ELAEDPAYAI  
1151 AHGVALNTDY YFSHLLGAAC VTFKALFGNN AKITESLLKR FIPEVWHPPD  
50 1201 DVAARLRAAG FGAVGAGATA EETRRMLHRA FDTLA\*

## SEQ.ID.NO. 9 DNA sequence of HSV polymerase gene for HSV1-DJL-M1

1 ATGTTTTCCG GTGGCGGCGG CCCGCTGTCC CCCGGAGGAA AGTCGGCGGC  
5 51 CAGGGCGGCG TCCGGGTTTT TTGCGCCCGC CGGCCCTCGC GGAGCCGGCC  
101 GGGGACCCCC GCCTTGTTTG AGGCAAACT TTTACAACCC CTACCTCGCC  
151 CCAGTCGGGA CGCAACAGAA GCCGACCGGG CCAACCCAGC GCCATACGTA  
10 201 CTATAGCGAA TGCGATGAAT TTCGATTCAT CGCCCCGCGG GTGCTGGACG  
251 AGGATGCCCC CCCGGAGAAG CGCGCCGGGG TGCACGACGG TCACCTCAAG  
15 301 CGCGCCCCCA AGGTGTACTG CGGGGGGGAC GAGCGCGACG TCCTCCGCGT  
351 CGGGTCGGGC GGCTTCTGGC CGCGGCGCTC GCGCCTGTGG GGCGGCGTGG  
401 ACCACGCCCC GGCGGGGTTC AACCCACCG TCACCGTCTT TCACGTGTAT  
20 451 GACATCCTGG AGAACGTGGA GCACGCGTAC GGCATGCGCG CGGCCAGTT  
501 CCACGCGCGG TTTATGGACG CCATCACACC GACGGGGACC GTCATCACGC  
25 551 TCCTGGGCCT GACTCCGGAA GGCCACCGGG TGGCCGTTCA CGTTTACGGC  
601 ACGCGGCAGT ACTTTTACAT GAACAAGGAG GAGGTTGACA GGCACCTACA  
651 ATGCCGCGCC CCACGAGATC TCTGCGAGCG CATGGCCGCG GCCCTGCGCG  
30 701 AGTCCCCGGG CGCGTCGTTT CGCGGCATCT CCGCGGACCA CTTCGAGGCG  
751 GAGGTGGTGG AGCGCACCGA CGTGTACTAC TACGAGACGC GCCCCGCTCT  
35 801 GTTTTACCGC GTCTACGTCC GAAGCGGGCG CGTGCTGTCT TACCTGTGCG  
851 ACAACTTCTG CCCGGCCATC AAGAAGTACG AGGGTGGGGT CGACGCCACC  
901 ACCCGGTTCA TCCTGGACAA CCCCAGGTTC GTCACCTTCG GCTGGTACCG  
40 951 TCTCAAACCG GGCCGGAACA ACACGCTAGC CCAGCCGCGG GCCCCGATGG  
1001 CCTTCGGGAC ATCCAGCGAT GTCGAGTTTA ACTGTACGGC GGACAACCTG  
45 1051 GCCATCGAGG GGGGCATGAG CGACCTACCG GCATACAAGC TCATGTGCTT  
1101 CGATATCGAA TGCAAGGCGG GGGGGGAGGA CGAGCTGGCC TTTCCGGTGG  
1151 CCGGGCACCC GGAGGACCTG GTCATCCAGA TATCCTGTCT GCTCTACGAC  
50 1201 CTGTCCACCA CCGCCCTGGA GCACGTCCTC CTGTTTTCGC TCGGTTCTTG  
1251 CGACCTCCCC GAATCCCACC TGAACGAGCT GGCGGCCAGG GGCCTGCCCCA  
55 1301 CGCCCGTGGT TCTGGAATTC GACAGCGAAT TCGAGATGCT GTTGGCCTTC  
1351 ATGACCCTTG TGAAACAGTA CGGCCCCGAG TTCGTGACCG GGTACAACAT  
1401 AATCAACTTC GACTGGCCCT TCTTGCTGGC CAAGCTGACG GACATTTACA

1451 AGGTCCCCCT GGACGGGTAC GGCCGCATGA ACGGCCGGGG CGTGTTTCGC  
1501 GTGTGGGACA TAGGCCAGAG CCACTTCCAG AAGCGCAGCA AGATAAAGGT  
5 1551 GAACGGCATG GTGAACATCG ACATGTACGG GATTATAACC GACAAGATCA  
1601 AGCTCTCGAG CTACAAGCTC AACGCCGTGG CCGAAGCCGT CCTGAAGGAC  
10 1651 AAGAAGAAGG ACCTGAGCTA TCGCGACATC CCCACCTACT ACGCCGCCGG  
1701 GCCCGCGCAA CGCGGGGTGA TCGGCGAGTA CTGCATACAG GATTCCCTGC  
1751 TGGTGGGCCA GCTGTTTTTT AAGTTTTTGC CCCATCTGGA GCTCTCGGCC  
15 1801 GTCGCGCGCT TGGCGGGTAT TAACATCACC CGCACCATCT ACGACGGCCA  
1851 GCAGATCCGC GTCTTTACGT GCCTGCTGCG CCTGGCCGAC CAGAAGGGCT  
20 1901 TTATTCTGCC GGACACCCAG GGGCGATTTA GGGGCGCCGG GGGGGAGGCG  
1951 CCCAAGCGTC CGGCCGCAGC CCGGGAGGAC GAGGAGCGGC CAGAGGAGGA  
2001 GGGGGAGGAC GAGAACGAAC GCGAGGAGGG CGGGGGCGAG CGGGAGCCGG  
25 2051 AGGGCGCGCG GGAGACCGCC GGCCGGCACG TGGGGTACCA GGGGGCCAGG  
2101 GTCCTTGACC CCACTTCCGG GTTTCACGTG AACCCCGTGG TGGTGTTCTGA  
30 2151 CTTTGCCAGC CTGTACCCCA GCATCATCCA GGCCCACAAC CTGTGCTTCA  
2201 GCACGCTCTC CCTGAGGGCC GACGCAGTGG CGCACCTGGA GGCGGGCAAG  
2251 GACTACCTGG AGATCGAGGT GGGGGGGCGA CGGCTGTTCT TCGTCAAGGC  
35 2301 TCACGTGCGA GAGAGCCTCC TCAGCATCCT CCTGCGGGAC TGGCTCGCCA  
2351 TGCGAAAGCA GATCCGCTCG CGGATTCCCC AGAGCAGCCC CGAGGAGGCC  
40 2401 GTGCTCCTGG ACAAGCAGCA GGCCGCCATC AAGGTCGTGT GTAACCTCGGT  
2451 TTACGGGTTC ACGGGAGCGC AGCACGGA CTGCGCGTGC CTGCACGTTG  
2501 CCGCGACGGT GACGACCATC GGCCGCGAGA TGCTGCTCGC GACCCGCGAG  
45 2551 TACGTCCACG CGCGCTGGGC GGCCTTCGAA CAGCTCCTGG CCGATTTCCT  
2601 GGAGGCGGCC GACATGCGCG CCCCCGGGCC CTATTCCATG CGCATCATCT  
50 2651 ACGGGGACAC GGACTCCATA TTTGTGCTGT GCCGCGCCT CACGGCCGCC  
2701 GGGCTGACGG CCGTGGGCGA CAAGATGGCG AGCCACATCT CGCGCGCGCT  
2751 GTTCTGCCC CCCATCAAAC TCGAGTGCGA AAAGACGTTT ACCAAGCTGC  
55 2801 TGCTGATCGC CAAGAAAAAG TACATCGGCG TCATCTACGG GGGTAAGATG  
2851 CTCATCAAGG GCGTGGATCT GGTGCGCAA ACAAAGTGG CGTTTATCAA  
60 2901 CCGCACCTCC AGGGCCCTGG TCGACCTGCT GTTTTACGAC GATACCGTAT

2951 CCGGAGCGGC CGCCGCGTTA GCCGAGCGCC CCGCAGAGGA GTGGCTGGCG  
3001 CGACCCCTGC CCGAGGGACT GCAGGCGTTC GGGGCCGTCC TCGTAGACGC  
5 3051 CCATCGGCGC ATCACCGACC CGGAGAGGGA CATCCAGGAC TTTGTTCTCA  
3101 CCGCCGA ACT GAGCAGACAC CCGCGCGCGT ACACCAACAA GCGCCTGGCC  
10 3151 CACCTGACGG TGTATTACAA GTCATGGCC CGCCGCGCGC AGGTCCCGTC  
3201 CATCAAGGAC CGGATCCCGT ACGTGATCGT GGCCCAGACC CGCGAGGTAG  
3251 AGGAGACGGT CGCGCGGCTG GCCGCCCTCC GCGAGCTAGA CGCCGCCGCC  
15 3301 CCAGGGGACG AGCCCGCCCC CCGCGGGCC CTGCCCTCCC CGGCCAAGCG  
3351 CCCCCGGGAG ACGCCGTCGC CTGCCGACCC CCCGGGAGGC GCGTCCAAGC  
20 3401 CCCGCAAGCT GCTGGTGTCC GAGCTGGCCG AGGATCCCGC ATACGCCATT  
3451 GCCCACGGCG TCGCCCTGAA CACGGA CTAT TACTTCTCCC ACCTGTTGGG  
3501 GCGGCGGTGC GTGACATTCA AGGCCCTGTT TGGGAATAAC GCCAAGATCA  
25 3551 CCGAGAGTCT GTTAAAAAGG TTTATTCCCG AAGTGTGGCA CCCCCGGAC  
3601 GACGTGGCCG CGCGGCTCCG GACCGCAGGG TTCGGGGCGG TGGGTGCCGG  
30 3651 CGCTACGGCG GAGGAACTC GTCGAATGTT GCATAGAGCC TTTGATACTC  
3701 TAGCATGA

## SEQ.ID.NO. 10 Amino acid sequence of DNA polymerase for HSV1-DJL-M1

1 MFSGGGGPLS PGGKSAARAA SGFFAPAGPR GAGRGPPPCL RQNFYNPYLA  
5 51 PVGTQQKPTG PTQRHTYYSE CDEFRFIAPR VLDEDAPEK RAGVHDGHLK  
101 RAPKVYCGGD ERDVLRVGSG GFWPRRSRLW GGVDHAPAGF NPTVTVFHVY  
151 DILENVEHAY GMRAAQFHAR FMDAITPTGT VITLLGLTPE GHRVAVHVY  
10 201 TRQYFYMNKE EVDRHLQCRA PRDLCERMAA ALRESPGASF RGISADHFEA  
251 EVVERTDVYY YETRPALFYR VYVRSGRVLS YLCDNFCPAI KKYEGGVDAT  
15 301 TRFILDNPGF VTFGWYRLKP GRNNTLAQPR APMAFGTSSD VEFNCTADNL  
351 AIEGGMSDLP AYKLMCFDIE CKAGGEDELA FPVAGHPEDL VIQISCLLYD  
401 LSTTALEHVL LFSLGSCDLP ESHLNELAAR GLPTPVVLEF DSEFEMLLAF  
20 451 MTLVKQYGPE FVTGYNIINF DWPFLAKLT DIYKVPLDGY GRMNGRGVFR  
501 VWDIGQSHFQ KRSEKIVNGM VNIDMYGIT DKIKLSSYKL NAVAEAVLKD  
25 551 KKKDLSYRDI PTYYAAGPAQ RGVIGEYCIQ DSLLVGQLFF KFLPHLELSA  
601 VARLAGINIT RTIYDGQQIR VFTCLRLAD QKGFIPLDTQ GRFRGAGGEA  
651 PKRPAAARED EERPEEEGED ENEREEGGGE REPEGARETA GRHVGYQGAR  
30 701 VLDPTSGFHV NPVVVFDFAS LYPYIIQAHN LCFSTLSLRA DAVAHLEAGK  
751 DYLEIEVGGR RLFFVKAHVR ESLLSILLRD WLAMRKQIRS RIPQSSPEEA  
35 801 VLLDKQAAI KVCNSVYGF TGAQHGLLPC LHVAATVTIT GREMLLATRE  
851 YVHARWAAFE QLLADFPEAA DMRAPGPYSM RIYGD TDSI FVLCRGLTAA  
901 GLTAVGDKMA SHISRALFLP PIKLECEKTF TKLLLIAKKK YIGVIYGGKM  
40 951 LIKGVDLVRK NNCAFINRTS RALVDLLFYD DTVSGAAAAL AERPAAEWLA  
1001 RPLPEGLQAF GAVLVDAHRR ITDPERDIQD FVLTAELSRH PRAYTNKRLA  
45 1051 HLTVYYKLMA RRAQVPSIKD RPYVIVAQT REVEETVARL AALRELDAAA  
1101 PGDEPAPPAA LPSPAKRPRE TPSPADPPGG ASKPRKLLVS ELAEDPAYAI  
1151 AHGVALNTDY YFSHLLGAAC VTFKALFGNN AKITESLLKR FIPEVWHPPD  
50 1201 DVAARLRTAG FGAVGAGATA EETRRMLHRA FDTLA\*

**SEQ.ID.NO. 11**      DNA sequence of DNA polymerase gene for HMCV-AD169-M1

1 ATGTTTTTCA ACCCGTATCT GAGCGGCGGC GTGACCGGCG GTGCGGTTCG  
5 51 GGGTGGCCGG CGTCAGCGTT CGCAGCCCGG CTCCGCGCAG GGCTCGGGCA  
101 AGCGGCCGCC ACAGAAACAG TTTTTCAGA TCGTGCCGCG AGGTGTCATG  
151 TTCGACGGTC AGACGGGGTT GATCAAGCAT AAGACGGGAC GGCTGCCTCT  
10 201 CATGTTCTAT CGAGAGATTA AACATTTGTT GAGTCATGAC ATGGTTTGGC  
251 CGTGTCTTG GCGCGAGACC CTGGTGGGTC GCGTGGTGGG ACCTATTCGT  
15 301 TTTCACACCT ACGATCAGAC GGACGCCGTG CTCTTCTTCG ACTCGCCCCA  
351 AAACGTGTCG CCGCGCTATC GTCAGCATCT GGTGCCTTCG GGAACGTGT  
401 TCGTTTTCTT CGGGGCCACA GAACACGGCT ACAGTATCTG CGTCAACGTT  
20 451 TTCGGGCAGC GCAGCTACTT TTA CTGTGAG TACAGCGACA CCGATAGGCT  
501 GCGTGAGGTC ATTGCCAGCG TGGGCGAACT AGTGCCCGAA CCGCGGACGC  
25 551 CATA CGCCGT GTCTGTCACG CCGGCCACCA AGACCTCCAT CTATGGGTAC  
601 GGGACGCGAC CCGTGCCCGA TTTGCAGTGT GTGTCTATCA GCAACTGGAC  
651 CATGGCCAGA AAAATCGGCG AGTATCTGCT GGAGCAGGGT TTTCCCGTGT  
30 701 ACGAGGTCCG TGTGGATCCG CTGACGCGTT TGGTCATCGA TCGGCGGATC  
751 ACCACGTTTCG GCTGGTGCTC CGTGAATCGT TACGACTGGC GGCAGCAGGG  
35 801 TCGCGCGTCG ACTTGTGATA TCGAGGTAGA CTGCGATGTC TCTGACCTGG  
851 TGGCTGTGCC CGACGACAGC TCGTGGCCGC GCTATCGATG CCTGTCTTC  
901 GATATCGAGT GCATGAGCGG CGAGGGTGGT TTTCCCTGCG CCGAGAAGTC  
40 951 CGATGACATT GTCATTCAGA TCTCGTGCGT GTGCTACGAG ACGGGGGGAA  
1001 ACACCGCCGT GGATCAGGGG ATCCCAAACG GGAACGATGG TCGGGGCTGC  
45 1051 ACTTCGGAGG GTGTGATCTT TGGGCACTCG GGTCTTCATC TCTTTACGAT  
1101 CGGCACCTGC GGGCAGGTGG GCCCAGACGT GGACGTCTAC GAGTTCCCTT  
1151 CCGAATACGA GCTGCTGCTG GGCTTTATGC TTTTCTTTCA ACGGTACGCG  
50 1201 CCGGCCTTTG TGACCGGTTA CAACATCAAC TCTTTTGA CT TGAAGTACAT  
1251 CCTCACGCGT CTCGAGTACC TGTATAAGGT GGA CTGCGAG CGCTTCTGCA  
55 1301 AGTTGCCTAC GCGCAGGGG GGCCGTTTCT TTTTACACAG CCCC GCCGTG  
1351 GGTTTTAAGC GGCAGTACGC CGCCGCTTTT CCCTCGGCTT CTCACAACAA  
1401 TCCGGCCAGC ACGGCCGCCA CCAAGGTGTA TATTGCGGGT TCGGTGGTTA

1451 TCGACATGTA CCCTGTATGC ATGGCCAAGA CTAACTCGCC CAACTATAAG  
1501 CTCAACACTA TGGCCGAGCT TTACCTGCGG CAACGCAAGG ATGACCTGTC  
5 1551 TTACAAGGAC ATCCCGCGTT GTTTCGTGGC TAATGCCGAG GGCCGCGCCC  
1601 AGGTAGGCCG TTA CTGTCTG CAGGACGCCG TATTGGTGCG CGATCTGTTC  
1651 AACACCATTA ATTTTCACTA CGAGGCCGGG GCCATCGCGC GGCTGGCTAA  
10 1701 AATTCGGTTG CGGCGTGTCA TCTTTGACGG ACAGCAGATC CGTATCTACA  
1751 CCTCGCTGCT GGACGAGTGC GCCTGCCGCG ATTTTATCCT GCCCAACCAC  
15 1801 TACAGCAAAG GTACGACGGT GCCCGAAACG AATAGCGTTG CTGTGTCACC  
1851 TAACGCTGCT ATCATCTCTA CCGCCGCTGT GCCCGGCGAC GCGGGTTCTG  
1901 TGGCGGCTAT GTTTCAGATG TCGCCGCCCT TGCAATCTGC GCCGTCCAGT  
1951 CAGGACGGCG TTTCACCCGG CTCCGGCAGT AACAGTAGTA GCAGCGTCGG  
2001 CGTTTTTCAGC GTCGGCTCCG GCAGTAGTGG CGGCGTCGGC GTTTCACAG  
25 2051 ACAATCACGG CGCCGGCGGT ACTGCGGCGG TTTCGTACCA GGGCGCCACG  
2101 GTGTTTGAGC CCGAGGTGGG TTA CTACAAC GACCCCGTGG CCGTGTTCGA  
2151 CTTTGCCAGC CTCTACCCCT CCATCATCAT GGCCACAAAC CTCTGCTACT  
2201 CCACCCTGCT GGTGCCGGGT GGCAGTACC CTGTGGACCC CGCCGACGTA  
2251 TACAGCGTCA CGCTAGAGAA CGGCGTGACC CACCGCTTTG TGCGTGCTTC  
35 2301 GGTGCGCGTC TCGGTGCTCT CGGAACTGCT CAACAAGTGG GTTTCGCAGC  
2351 GCGTGCCGT GCGCGAATGC ATGCGCGAGT GTCAAGACCC TGTGCGCCGT  
2401 ATGCTGCTCG ACAAGGAACA GATGGCGCTC AAAGTAACGT GCAACGCTTT  
2451 CTACGGTTTT ACCGGCGCGC TGAACGGTAT GATGCCGTGT CTGCCATCG  
2501 CCGCCAGCAT CACGCGCATC GGTCGCGACA TGCTAGAGCG CACGGCGCGG  
45 2551 TTCATCAAAG ACAACTTTTC AGAGCCGTGT TTTTGCACA ATTTTTTTAA  
2601 TCAGGAAGAC TATGTAGTGG GAACGCGGGA GGGGGATTCTG GAGGAGAGCA  
2651 GCGCGTTACC GGAGGGGCTC GAAACATCGT CAGGGGGCTC GAACGAACGG  
50 2701 CGGGTGGAGG CGCGGGTCAT CTACGGGGAC ACGGACAGCG TGTTTGTCCG  
2751 CTTTCGTGGC CTGACGCCGC AGGCTCTGGT GCGCGTGGG CCCAGCCTGG  
55 2801 CGCACTACGT GACGGCCTGT CTTTTTGTGG AGCCCGTCAA GCTGGAGTTT  
2851 GAAAAGGTCT TCGTCTCTCT TATGATGATC TGCAAGAAAC GTTACATCGG  
60 2901 CAAAGTGGAG GCGCCTCGG GTCTGAGCAT GAAGGGCGTG GATCTGGTGC

2951 GCAAGACGGC CTGCGAGTTC GTCAAGGGCG TCACGCGTGA CGTCCTCTCG  
3001 CTGCTCTTTG AGGATCGCGA GGTCTCGGAA GCAGCCGTGC GCCTGTGCGG  
5 3051 CCTCTCACTC GATGAAGTCA AGAAGTACGG CGTGCCACGC GGTTTCTGGC  
3101 GTATCTTACG CCGCTTGGTG CAGGCCCCGCG ACGATCTGTA CCTGCACCGT  
10 3151 GTGCGTGTCG AGGACCTGGT GCTTTCGTCG GTGCTCTCTA AGGACATCTC  
3201 GCTGTACCGT CAATCTAACC TGCCGCACAT TGCCGTCATT AAGCGATTGG  
3251 CGGCCCCGTT TGAGGAGCTA CCCTCGGTCG GGGATCGGGT CTTTTACGTT  
15 3301 CTGACGGCGC CCGGTGTCCG GACGGCGCCG CAGGGTTCCT CCGACAACGG  
3351 TGATTCTGTA ACCGCCGGCG TGGTTTCCCG GTCGGACGCG ATTGATGGCA  
20 3401 CGGACGACGA CGCTGACGGC GGCAGGGTAG AGGAGAGCAA CAGGAGAGGA  
3451 GGAGAGCCGG CAAAGAAGAG GGCGCGGAAA CCACCGTCGG CCGTGTGCAA  
3501 CTACGAGGTA GCCGAAGATC CGAGCTACGT GCGCGAGCAC GCGTGCCCA  
25 3551 TTCACGCCGA CAAGTACTTT GAGCAGGTTT TCAAGGCTGT AACTAACGTG  
3601 CTGTCGCCCC TCTTCCCCG CGGCGAAACC GCGCGCAAAG ACAAGTTTTT  
30 3651 GCACATGGTG CTGCCGCGGC GCTTGCACTT GGAGCCGGCT TTTCTGCCGT  
3701 ACAGTGTC AA GGCGCACGAA TGCTGTTGA

**SEQ. ID. NO. 12      Amino acid sequence of DNA polymerase for HCMV-AD169-M1**

1 MFFNPYLSGG VTGGAVAGGR RQRSQPGSAQ GSGKRPPQKQ FLQIVPRGVM  
 5 51 FDGQTGLIKH KTGRPLPMFY REIKHLLSHD MYWPCPWRET LVGRVVGP  
 101 FHTYDQTDV LFFDSPENV PRYRQHLVPS GNVLRFFGAT EHGYSICVNV  
 10 151 FGQRSYFYCE YSDTDRLREV IASVGELVPE PRTPYAVSVT PATKTSIYGY  
 201 GTRPVPDLQC VSISNWTMAR KIGEYLLSEQ FPVYEVVRDP LTRLVIDRRI  
 251 TTFGWCSVNR YDWRQQGRAS TCDIEVDCDV SDLVAVPDDS SWPRYRCLSF  
 15 301 DIECMSGEGG FPCAESDDI VIQISVCYE TGGNTAVDQG IPNGNDGRGC  
 351 TSEGVIHGHS GLHLFTIGTC GQVGPDVDVY EFPSEYELL GFMLFFQRYA  
 401 PAFVTGYNIN SFDLKYILTR LEYLYKVDSQ RFCKLPTAQG GRFFLHSPAV  
 451 GFKRQYAAAF PSASHNNPAS TAATKVYIAG SVVIDMYPVC MAKTNPNYK  
 501 LNTMAELYLR QRKDDLSYKD IPRCFVANAE GRAQVGRYCL QDAVLVRDLF  
 25 551 NTINFHYEAG AIARLAKIPL RRVIFDGQOI RIYTSLLDEC ACRDFILPNH  
 601 YSKGTTVPET NSVAVSPNAA IISTAAVPGD AGSVAAMFQM SPPLQSAPSS  
 30 651 QDGVSPGSGS NSSSSVGVFS VSGSSGGVG VSNDNHGAGG TAAVSYQGAT  
 701 VFEPEVGYYN DPVAVFDFAS LYPSTMAHN LCYSTLLVPG GEYPVDPADV  
 751 YSVTLENGVT HRFVRASVRV SVLSELLNKW VSQRAVREC MRECQDPVRR  
 35 801 MLLDKEQMAL KVTCTAFYGF TGALNGMMPC LPIAASITRI GRDMLERTAR  
 851 FIKDNFSEPC FLHNFFNQED YVVGTTREGDS EESSALPEGL ETSSGGSNER  
 40 901 RVEARVTYGD TDSVFVRFRG LTPQALVARG PSLAHYVTAC LFVEPVKLEF  
 951 EKVFSVSLMMI CKKRYIGKVE GASGLSMKGV DLVRKTACEF VKGVTRDVL  
 1001 LLFEDREVSE AAVRLSRLSL DEVKKYGVPR GFWRLRLRV QARDDLHLR  
 45 1051 VRVEDLVLS VLSKDISLYR QSNLPHIAVI KRLAARSEEL PSVGDRVFYV  
 1101 LTAPGVRTAP QGSSDNGDSV TAGVVSRS DA IDGTDDADG GGVEESNRRG  
 50 1151 GEPAKKRARK PPSAVCNIEV AEDPSYVREH GVPIHADKYF EQVLKAVTNV  
 1201 LSPVFPGET ARKDKFLHMV LPRRLHLEPA FLPYSVKAHE CC\*

Figure 6

SEQ.ID.NO.13

Amino acid sequence of DNA polymerase for HCMV-AD169

5     1 MFFNPYLSGG VTGGAVAGGR RQRSQPGSAQ GSGKRPPQKQ FLQIVPRGVM  
       51 FDGQTGLIKH KTGRPLPMFY REIKHLLSHD MVWPCPWRET LVGRVVGPRI  
       101 FHTYDQTDVAV LFFDSPENVV PRYRQHLVPS GNVLRFFGAT EHGYSICVNV  
 10     151 FGQRSYFYCE YSDTDLREV IASVGELVPE PRTPYAVSVT PATKTSIYGY  
       201 GTRPVLDLQC VSISNWTMAR KIGEYLLQEG FPVYEVVRDP LTRLVIDRRI  
       251 TTFGWCSVNR YDWRQQGRAS TCDIEVDCDV SDLVAVPDDS SWPRYRCLSF  
 15     301 DIECMSGEGG FPCAESDDI VIQISCVCYE TGGNTAVDQG IPNGNDGRGC  
       351 TSEGVIFGHS GLHLFTIGTC GQVGPDVDVY EFPSEYELLL GFMLFFQRYA  
 20     401 PAFVTGYNIN SFDLKYILTR LEYLYKVDSQ RFCKLPTAQQ GRFFLHSPAV  
       451 GFKRQYAAAF PSASHNNPAS TAATKVYIAG SVVIDMYPVC MAKTNPNYK  
       501 LNTMAELYLR QRKDDLSYKD IPRCFVANAE GRAQVGRYCL QDAVLVRDLF  
 25     551 NTINFHYEAG AIARLAKIPL RRVIFDGQOI RIYTSLLDEC ACRDFILPNH  
       601 YSKGTTVPET NSVAVSPNAA IISTAAVPGD AGSVAAMFQM SPPLQSAPSS  
 30     651 QDGVSPGSGS NSSSSVGVFS VGSGSSGGVG VSNDNHGAGG TAAVSYQGAT  
       701 VFEPEVGYYN DPVAVVDFAS LYPSTMAHN LCYSTLLVPG GEYPVDPADV  
       751 YSVTLENGVT HRFVRASVRV SVLSELLNKW VSQRRVREC MRECQDPVRR  
 35     801 MLLDKEQMAL KVTCLAFYGF TGVVNGMMPC LPIAASITRI GRDMLERTAR  
       851 FIKDNFSEPC FLHNFFNQED YVVGTTREGDS EESSALPEGL ETSSGGSNER  
 40     901 RVEARVIYGD TDSVFVRFRG LTPQALVARG PSLAHYVTAC LFVEPVKLEF  
       951 EKVFSVSLMMI CKKRYIGKVE GASGLSMKGV DLVRKTACEF VKGVTRDVLV  
       1001 LLFEDREVSE AAVRLSRLSL DEVKKYGVPR GFWRLRLRV QARDDLHLHR  
       1051 VRVEDLVLSV VLSKDISLYR QSNLPHIAVI KRLAARSEEL PSVGDRVFYV  
       1101 LTAPGVRTAP QGSSDNGDSV TAGVVSRSDA IDGTDDDDADG GGVEESNRRG  
 50     1151 GEPAKKRARK PPSAVCNIEV AEDPSYVREH GVPIHADKYF EQVLKAVTNV  
       1201 LSPVFPGET ARKDKFLHNV LPRRLHLEPA FLPYSVKAHE CC\*  
 55

## SEQUENCE LISTING

<110> Homa, Fred  
 Wathen, Michael  
 Hopkins, Todd  
 Thomsen, Darrell

<120> A Method for Treating Herpes Virus

<130> 00221

<160> 19

<170> PatentIn version 3.0

<210> 1

<211> 3717

<212> DNA

<213> herpes simplex

<400> 1

```

atgttttgtg cgcgggcgcg cccgacttcc cccgggggga agtcggcggc tcgggcggcg      60
tctgggtttt ttgccccca caaccccgcg ggagccaccc agacggcacc gccgccttgc      120
cgccggcaga acttctacaa cccccacctc gctcagaccg gaacgcagcc aaaggcccc      180
gggcccggctc agcgccatac gtactacagc gagtgcgacg aatttcgatt tatcgccccg      240
cgttcgctgg acgaggacgc ccccgcgag cagcgcaccg ggggtccacga cggccgcctc      300
cggcgcgccc ctaaggtgta ctgcgggggg gacgagcgcg acgtcctccg cgtgggcccc      360
gagggcttct ggccgcgtcg cttgcgcctg tggggcggtg cggaccatgc cccaagggg      420
ttcgacccca ccgtcacctg cttccacgtg tacgacatcc tggagcacgt ggaacacgcg      480
tacagcatgc gcgccgcca gctccacgag cgatttatgg acgccatcac gccgcgggg      540
accgtcatca cgcttctggg tctgaccccc gaaggccatc gcgtcgccgt tcacgtctac      600
ggcacgcggc agtactttta catgaacaag gcggaggtgg atcggcacct gcagtgcctg      660
gccccgcgcg atctctgcga gcgcctggcg gcggccctgc gcgagtcgcc gggggcgctg      720
ttccgcggca tctccgcgga ccacttcgag gcggaggtgg tggagcgcg cgcgtgtac      780
tattacgaaa cgcgcccagc cctgtactac cgcgtcttcg tgcgaagcgg gcgcgcgctg      840
gcctacctgt gcgacaactt ttgccccgcg atcaggaagt acgagggggg cgtcgacgcc      900
accacccggt ttatcctgga caaccggggg tttgtcacct tcggctggta ccgcctcaag      960
cccggccgcg ggaacgcgcc ggcccaaccg cccccccga cggcgttcgg aacctcgagc     1020
gacgtcgagt ttaactgcac ggcggacaac ctggccgctc agggggccat gtgtgacctg     1080
ccggcctaca agtcatgtg cttcgatata gaatgcaagg ccggggggga ggacgagctg     1140
gcctttccgg tcgcggaacg cccggaagac ctcgtcatcc agatctcctg tctgtcttac     1200
gacctgtcca ccaccgcct cgagcacatc ctctgtttt cgctcggatc ctgcgacctc     1260

```

cccgagtccc acctcagcga tctcgccctcc agggggcctgc cggccccccgt cgtcctggag 1320  
 tttgacagcg aattcgagat gctgctggcc ttcacgacct tcgtcaagca gtacggcccc 1380  
 gagttcgtga cggggtacaa catcatcaac ttcgactggc ccttcgtcct gaccaagctg 1440  
 acggagatct acaagggtccc gctcgacggg tacggggcgca tgaacggccg ggggtgtgttc 1500  
 cgcgtgtggg acatcggcca gagccacttt cagaagcgca gcaagatcaa ggtgaacggg 1560  
 atgggtgaaca tcgacatgta cggcatcatc accgacaagg tcaaactctc cagctacaag 1620  
 ctgaacgccg tcgccgaggc cgtcttgaag gacaagaaga aggatctgag ctaccgcgac 1680  
 atccccgcct actacgcctc cggggcccgcg cagcgccggg tgatcggcga gtattgtgtg 1740  
 caggactcgc tgctggtcgg gcagctgttc ttcaagtttc tgccgcacct ggagctttcc 1800  
 gccgtcgcgc gcctggcggg catcaacatc acccgcacca tctacgacgg ccagcagatc 1860  
 cgcgtcttca cgtgcctcct gcgccttgcg ggccagaagg gcttcacact gccggacacc 1920  
 cagggggcgt ttcggggcct cgacaaggag gcgcccgaag gcccggccgt gcctcggggg 1980  
 gaaggggagc ggccggggga cgggaacggg gacgaggata aggacgacga cgaggacgag 2040  
 gacggggacg agcgcgagga ggtcgcgcgc gagaccgggg gccggcacgt tgggtaccag 2100  
 gggggccggg tcctcgaccc cacctccggg ttacacgtcg acccgtggg ggtgtttgac 2160  
 tttgccagcc tgtacccag catcatccag gccacaacc tgtgcttcag tacgctctcc 2220  
 ctgcggcccc aggccgtcgc gcacctggag gcggaccggg actacctgga gatcgaggtg 2280  
 gggggccgac ggctgttctt cgtgaaggcc cacgtacgcg agagcctgct gagcatcctg 2340  
 ctgcgcgact ggctggccat gcgaaagcag atccgctcgc ggatccccca gagcaccccc 2400  
 gaggaggccg tcctcctcga caagcaacag gccgccatca aggtggtgtg caactcgggtg 2460  
 tacgggttca cggggcgca gcacggtctt ctgccctgcc tgcaagtggt cgccaccgtg 2520  
 acgaccatcg gccgcgagat gctcctcgcg acgcgcgcgt acgtgcacgc gcgctgggcg 2580  
 gagttcgatc agctgctggc cgactttccg gaggcggccg gcatgcgcgc ccccgggtccg 2640  
 tactccatgc gcatcatcta cggggacacg gactccattt tcgttttgtg ccgcggcctc 2700  
 acggccgcgg gcctggtggc catgggcgac aagatggcga gccacatctc gcgcgcgctg 2760  
 ttcctcccc cgatcaagct cgagtgcgaa aaaacgttca ccaagctgct gctcatcgcc 2820  
 aagaaaaagt acatcggcgt catctgcggg ggcaagatgc tcatcaaggg cgtggatctg 2880  
 gtgcgcaaaa acaactgcgc gtttatcaac cgacactcca gggccctggg cgacctgctg 2940  
 ttttacgacg ataccgtatc cggagcggcc gccgcgttag ccgagcgccc cgcagaggag 3000  
 tggctggcgc gaccctgcc cgagggactg caggcgctcg gggccgtcct cgtagacgcc 3060  
 catcggcgca tcaccgacct ggagagggac atccaggact ttgtcctcac cgccgaactg 3120

```

agcagacacc cgcgcgcgta caccaacaag cgcctggccc acctgacggt gtattacaag 3180
ctcatggccc gccgcgcgca ggtcccgtcc atcaaggacc ggatcccgta cgtgatcgtg 3240
gcccagaccc gcgaggtaga ggagacggtc gcgcggctgg ccgccctccg cgagctagac 3300
gccgcgcgcc caggggacga gcccgcctcc ccagcggccc tgcctctccc ggccaagcgc 3360
ccccgggaga cgccgtcgca tgccgacccc ccgggaggcg cgtccaagcc ccgcaagctg 3420
ctggtgtccg agctggcgga ggatcccggg tacgccatcg cccggggcgt tccgctcaac 3480
acggactatt acttctcgca cctgctgggg gcggcctgcg tgacgttcaa ggccctgttt 3540
ggaaataacg ccaagatcac cgagagtctg ttaaagaggt ttattcccga gacgtggcac 3600
cccccgacg acgtggccgc gcggctcagg gccgcggggg tcgggccggc gggggccggc 3660
gctacggcgg aggaaactcg tcgaatgttg catagagcct ttgatactct agcatga 3717

```

```

<210> 2
<211> 1238
<212> PRT
<213> herpes simplex

```

```

<400> 2

```

```

Met Phe Cys Ala Ala Gly Gly Pro Thr Ser Pro Gly Gly Lys Ser Ala
1           5           10           15
Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro His Asn Pro Arg Gly Ala
20          25          30
Thr Gln Thr Ala Pro Pro Pro Cys Arg Arg Gln Asn Phe Tyr Asn Pro
35          40          45
His Leu Ala Gln Thr Gly Thr Gln Pro Lys Ala Pro Gly Pro Ala Gln
50          55          60
Arg His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro
65          70          75          80
Arg Ser Leu Asp Glu Asp Ala Pro Ala Glu Gln Arg Thr Gly Val His
85          90          95
Asp Gly Arg Leu Arg Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu
100         105         110
Arg Asp Val Leu Arg Val Gly Pro Glu Gly Phe Trp Pro Arg Arg Leu
115         120         125
Arg Leu Trp Gly Gly Ala Asp His Ala Pro Lys Gly Phe Asp Pro Thr
130         135         140
Val Thr Val Phe His Val Tyr Asp Ile Leu Glu His Val Glu His Ala
145         150         155         160
Tyr Ser Met Arg Ala Ala Gln Leu His Glu Arg Phe Met Asp Ala Ile
165         170         175

```

Thr Pro Ala Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly  
 180 185 190  
 His Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met  
 195 200 205  
 Asn Lys Ala Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp  
 210 215 220  
 Leu Cys Glu Arg Leu Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser  
 225 230 235 240  
 Phe Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg  
 245 250 255  
 Ala Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Thr Leu Tyr Tyr Arg Val  
 260 265 270  
 Phe Val Arg Ser Gly Arg Ala Leu Ala Tyr Leu Cys Asp Asn Phe Cys  
 275 280 285  
 Pro Ala Ile Arg Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe  
 290 295 300  
 Ile Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys  
 305 310 315 320  
 Pro Gly Arg Gly Asn Ala Pro Ala Gln Pro Arg Pro Pro Thr Ala Phe  
 325 330 335  
 Gly Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala  
 340 345 350  
 Val Glu Gly Ala Met Cys Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe  
 355 360 365  
 Asp Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val  
 370 375 380  
 Ala Glu Arg Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr  
 385 390 395 400  
 Asp Leu Ser Thr Thr Ala Leu Glu His Ile Leu Leu Phe Ser Leu Gly  
 405 410 415  
 Ser Cys Asp Leu Pro Glu Ser His Leu Ser Asp Leu Ala Ser Arg Gly  
 420 425 430  
 Leu Pro Ala Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu  
 435 440 445  
 Leu Ala Phe Met Thr Phe Val Lys Gln Tyr Gly Pro Glu Phe Val Thr  
 450 455 460  
 Gly Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Val Leu Thr Lys Leu  
 465 470 475 480  
 Thr Glu Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly  
 485 490 495  
 Arg Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys  
 500 505 510

Arg Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly  
 515 520 525  
 Ile Ile Thr Asp Lys Val Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val  
 530 535 540  
 Ala Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp  
 545 550 555 560  
 Ile Pro Ala Tyr Tyr Ala Ser Gly Pro Ala Gln Arg Gly Val Ile Gly  
 565 570 575  
 Glu Tyr Cys Val Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys  
 580 585 590  
 Phe Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile  
 595 600 605  
 Asn Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr  
 610 615 620  
 Cys Leu Leu Arg Leu Ala Gly Gln Lys Gly Phe Ile Leu Pro Asp Thr  
 625 630 635 640  
 Gln Gly Arg Phe Arg Gly Leu Asp Lys Glu Ala Pro Lys Arg Pro Ala  
 645 650 655  
 Val Pro Arg Gly Glu Gly Glu Arg Pro Gly Asp Gly Asn Gly Asp Glu  
 660 665 670  
 Asp Lys Asp Asp Asp Glu Asp Glu Asp Gly Asp Glu Arg Glu Glu Val  
 675 680 685  
 Ala Arg Glu Thr Gly Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val  
 690 695 700  
 Leu Asp Pro Thr Ser Gly Phe His Val Asp Pro Val Val Val Phe Asp  
 705 710 715 720  
 Phe Ala Ser Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe  
 725 730 735  
 Ser Thr Leu Ser Leu Arg Pro Glu Ala Val Ala His Leu Glu Ala Asp  
 740 745 750  
 Arg Asp Tyr Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val  
 755 760 765  
 Lys Ala His Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp  
 770 775 780  
 Leu Ala Met Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Thr Pro  
 785 790 795 800  
 Glu Glu Ala Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val  
 805 810 815  
 Cys Asn Ser Val Tyr Gly Phe Thr Gly Ala Gln His Gly Leu Leu Pro  
 820 825 830  
 Cys Leu His Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu

835	840	845
Leu Ala Thr Arg Ala Tyr Val His Ala Arg Trp Ala Glu Phe Asp Gln 850 855 860		
Leu Leu Ala Asp Phe Pro Glu Ala Ala Gly Met Arg Ala Pro Gly Pro 865 870 875 880		
Tyr Ser Met Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu 885 890 895		
Cys Arg Gly Leu Thr Ala Ala Gly Leu Val Ala Met Gly Asp Lys Met 900 905 910		
Ala Ser His Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu 915 920 925		
Cys Glu Lys Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr 930 935 940		
Ile Gly Val Ile Cys Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu 945 950 955 960		
Val Arg Lys Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu 965 970 975		
Val Asp Leu Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala 980 985 990		
Leu Ala Glu Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu 995 1000 1005		
Gly Leu Gln Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg 1010 1015 1020		
Ile Thr Asp Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala 1025 1030 1035		
Glu Leu Ser Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala 1040 1045 1050		
His Leu Thr Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val 1055 1060 1065		
Pro Ser Ile Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr 1070 1075 1080		
Arg Glu Val Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu 1085 1090 1095		
Leu Asp Ala Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala 1100 1105 1110		
Leu Pro Ser Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser His Ala 1115 1120 1125		
Asp Pro Pro Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser 1130 1135 1140		
Glu Leu Ala Glu Asp Pro Gly Tyr Ala Ile Ala Arg Gly Val Pro 1145 1150 1155		

Leu Asn Thr Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys  
 1160 1165 1170

Val Thr Phe Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu  
 1175 1180 1185

Ser Leu Leu Lys Arg Phe Ile Pro Glu Thr Trp His Pro Pro Asp  
 1190 1195 1200

Asp Val Ala Ala Arg Leu Arg Ala Ala Gly Phe Gly Pro Ala Gly  
 1205 1210 1215

Ala Gly Ala Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala  
 1220 1225 1230

Phe Asp Thr Leu Ala  
 1235

<210> 3  
 <211> 3723  
 <212> DNA  
 <213> herpes simplex

<400> 3  
 atgttttgtg ccgcgggagg cccggcttcc cccgggggga agtcggcggc tcgggcggcg 60  
 tctgggtttt ttgccccca caaccccgagg gagccaccc agacggcacc gccgccttgc 120  
 cgccggcaga acttctacaa cccccacctc gctcagaccg gaacgcagcc aaaggccccc 180  
 gggccggctc agcgccatac gtactacagc gagtgcgacg aatttcgatt tatcgccccg 240  
 cgttcgctgg acgaggacgc cccgcgggag cagcgccaccg gggtcacga cgccgcctc 300  
 cggcgcgccc ctaaggtgta ctgcgggggg gacgagcgcg acgtcctccg cgtgggccccg 360  
 gagggcttct ggccgcgtcg attgcgcctg tggggcggtg cggaccatgc ccccgagggg 420  
 ttcgacccca ccgtcaccgt cttccacgtg tacgacatcc tggagcacgt ggaacacgcg 480  
 tacagcatgc gcgcgcacca gctccacgag cgatttatgg acgccatcac gcccgccggg 540  
 accgtcatca cgcttctggg tctgaccccc gaaggccatc gcgtcgccgt tcacgtctac 600  
 ggcacgcggc agtactttta catgaacaag gcggaggtgg atcggcacct gcagtgccgt 660  
 gccccgcgcg atctctgcga gcgcctggcg gcggccctgc gcgagtcgcc gggggcgctc 720  
 ttccgcggca tctccgcgga ccacttcgag gcggaggtgg tggagcgcg cgacgtgtac 780  
 tattacgaaa cgcgccccgac cctgtactac cgcgtcttcg tgcgaagcgg gcgcgcgctg 840  
 gcctacctgt gcgacaactt ttgccccgag atcaggaagt acgagggggg cgtcgacgcc 900  
 accacccggt ttatcctgga caacccgggg tttgtcacct tcggctggta ccgcctcaag 960  
 cccggccgcg ggaacgcgcc ggcccaaccg cgccccccga cggcgcttcg aacctcgagc 1020  
 gacgtcgagt ttaactgcac ggcggacaac ctggccgctc agggggccat gtgtgacctg 1080  
 ccggcctaca agctcatgtg cttcgatatc gaatgcaagg ccggggggga ggacgagctg 1140

gcctttccgg tcgcggaacg cccggaagac ctcgtcatcc agatctcctg tctgctctac 1200  
 gacctgtcca ccaccgccct cgagcacatc ctctgtttt cgctcggatc ctgcgacctc 1260  
 cccgagtccc acctcagcga tctcgctccc aggggacctg cgccccccgt cgtcctggag 1320  
 tttgacagcg aattcgagat gctgctggcc ttcattgacct tcgtcaagca gtacggcccc 1380  
 gagttcgtga cggggtacaa catcatcaac ttcgactggc ccttcgtcct gaccaagctg 1440  
 acggagatct acaagggtccc gctcgacggg tacgggcgca tgaacggccg ggggtgtgttc 1500  
 cgcgtgtggg acatcggcca gagccacttt cagaagcgca gcaagatcaa ggtgaacggg 1560  
 atgggtgaaca tcgacatgta cggcatcatc accgacaagg tcaaactctc cagctacaag 1620  
 ctgaacgccg tcgccgaggc cgtcttgaag gacaagaaga aggatctgag ctaccgcgac 1680  
 atccccgcct actacgcctc cgggccccgc cagcgcgggg tgatcggcga gtattgtgtg 1740  
 caggactcgc tgctggtcgg gcagctgttc ttcaagtttc tgccgcacct ggagctttcc 1800  
 gccgtcgcgc gcctggcggg catcaacatc acccgacca tctacgacgg ccagcagatc 1860  
 cgcgtcttca cgtgcctcct gcgccttgcg ggccagaagg gcttcatcct gccggacacc 1920  
 cagggggcgt ttcggggcct cgacaaggag gcgccaagc gcccgccgt gcctcggggg 1980  
 gaaggggagc ggccggggga cgggaacggg gacgaggata aggacgacga cgaggacggg 2040  
 gacgaggacg gggacgagcg cgaggaggtc gcgcgcgaga ccgggggccc gcacgttggg 2100  
 taccaggggg cccgggtcct cgaccccacc tccgggtttc acgtcgacc cgtgggtggtg 2160  
 tttgactttg ccagcctgta cccagcatc atccaggccc acaacctgtg cttcagtagc 2220  
 ctctccctgc ggcccgaggc cgtcgcgcac ctggaggcgg accgggacta cctggagatc 2280  
 gaggtggggg gccgacggct gttcttcgtg aaggcccacg tacgcgagag cctgctgagc 2340  
 atcctgctgc gcgactggct ggccatgcga aagcagatcc gctcgcggat ccccagagc 2400  
 cccccgagg aggcgcctcct cctcgacaag caacaggccg ccatcaaggt ggtgtgcaac 2460  
 tcggtgtacg ggttcaccgg ggcgcagcac ggtcttctgc cctgcctgca cgtggccgcc 2520  
 accgtgacga ccatcggccg cgagatgctc ctgcgcagc gcgcgtacgt gcacgcgcgc 2580  
 tggggggagt tcgatcagct gctggccgac tttccggagg cggccggcat gcgcgcccc 2640  
 ggtccgtact ccatgcgcat catctacggg gacacggact ccattttcgt tttgtgcgc 2700  
 ggcttcacgg ccgcgggcct ggtggccatg ggcgacaaga tggcgagcca catctcgcgc 2760  
 gcgctgttcc tccccccgat caagctcgag tgcgaaaaaa cgttcaccaa gctgctgctc 2820  
 atcgccaaga aaaagtacat cggcgtcatc tgcgggggca agatgctcat caagggcgtg 2880  
 gatctggtgc gcaaaaacaa ctgcgcgttt atcaaccgca cctccagggc cctggtcgac 2940  
 ctgctgtttt acgacgatac cgtatccgga gcggccgccg cgtagccga gcgccccgca 3000

gaggagtggc tggcgcgacc cctgcccagag ggactgcagg cgttcggggc cgtcctcgta 3060  
 gacgcccata ggcgcatacac cgacccggag agggacatcc aggactttgt cctcaccgcc 3120  
 gaactgagca gacaccgcgc cgcgtacacc aacaagcgcc tggcccacct gacggtgtat 3180  
 tacaagctca tggcccgcgc cgcgcaggtc cgtccatca aggaccggat cccgtacgtg 3240  
 atcgtggccc agaccgcga ggtagaggag acggtcgcgc ggctggccgc cctccgcgag 3300  
 ctagacgcgc cgcgccagag ggacgagccc gccccccag cggccctgcc ctccccggcc 3360  
 aagcgcccc gggagacgcc gtcgcatgcc gacccccgg gaggcgcgtc caagccccgc 3420  
 aagctgctgg tgtccgagct ggcggaggat cccgggtacg ccatcgccgc gggcggttccg 3480  
 ctcaacacgg actattactt ctgcacctg ctgggggcgc cctgcgtgac gttcaaggcc 3540  
 ctgtttggaa ataacgcaa gatcaccgag agtctgttaa agaggtttat tcccagagacg 3600  
 tggcaccccc cggacgacgt ggccgcgcgc ctcagggccg cgggggttcgg gccgycgggg 3660  
 gccgycgcta cggcggagga aactcgtcga atgttgcata gagcctttga tactctagca 3720  
 tga 3723

<210> 4  
 <211> 1240  
 <212> PRT  
 <213> herpes simplex

<400> 4

Met Phe Cys Ala Ala Gly Gly Pro Ala Ser Pro Gly Gly Lys Ser Ala  
 1 5 10 15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro His Asn Pro Arg Gly Ala  
 20 25 30  
 Thr Gln Thr Ala Pro Pro Pro Cys Arg Arg Gln Asn Phe Tyr Asn Pro  
 35 40 45  
 His Leu Ala Gln Thr Gly Thr Gln Pro Lys Ala Pro Gly Pro Ala Gln  
 50 55 60  
 Arg His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro  
 65 70 75 80  
 Arg Ser Leu Asp Glu Asp Ala Pro Ala Glu Gln Arg Thr Gly Val His  
 85 90 95  
 Asp Gly Arg Leu Arg Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu  
 100 105 110  
 Arg Asp Val Leu Arg Val Gly Pro Glu Gly Phe Trp Pro Arg Arg Leu  
 115 120 125  
 Arg Leu Trp Gly Gly Ala Asp His Ala Pro Glu Gly Phe Asp Pro Thr  
 130 135 140  
 Val Thr Val Phe His Val Tyr Asp Ile Leu Glu His Val Glu His Ala

145		150		155		160
Tyr Ser Met Arg Ala	Ala Ala Gln Leu His	Glu Arg Phe Met Asp	Ala Ile			
	165	170	175			
Thr Pro Ala Gly Thr	Val Ile Thr Leu Leu Gly	Leu Thr Pro Glu Gly				
	180	185	190			
His Arg Val Ala Val	His Val Tyr Gly Thr	Arg Gln Tyr Phe Tyr Met				
	195	200	205			
Asn Lys Ala Glu Val	Asp Arg His Leu Gln Cys	Arg Ala Pro Arg Asp				
	210	215	220			
Leu Cys Glu Arg Leu	Ala Ala Ala Leu Arg	Glu Ser Pro Gly Ala Ser				
	225	230	235			240
Phe Arg Gly Ile Ser	Ala Asp His Phe Glu	Ala Glu Val Val Glu Arg				
	245	250	255			
Ala Asp Val Tyr Tyr	Tyr Glu Thr Arg Pro	Thr Leu Tyr Tyr Arg Val				
	260	265	270			
Phe Val Arg Ser Gly	Arg Ala Leu Ala Tyr	Leu Cys Asp Asn Phe Cys				
	275	280	285			
Pro Ala Ile Arg Lys	Tyr Glu Gly Gly Val	Asp Ala Thr Thr Arg Phe				
	290	295	300			
Ile Leu Asp Asn Pro	Gly Phe Val Thr Phe	Gly Trp Tyr Arg Leu Lys				
	305	310	315			320
Pro Gly Arg Gly Asn	Ala Pro Ala Gln Pro	Arg Pro Pro Thr Ala Phe				
	325	330	335			
Gly Thr Ser Ser Asp	Val Glu Phe Asn Cys	Thr Ala Asp Asn Leu Ala				
	340	345	350			
Val Glu Gly Ala Met	Cys Asp Leu Pro Ala	Tyr Lys Leu Met Cys Phe				
	355	360	365			
Asp Ile Glu Cys Lys	Ala Gly Gly Glu Asp	Glu Leu Ala Phe Pro Val				
	370	375	380			
Ala Glu Arg Pro Glu	Asp Leu Val Ile Gln	Ile Ser Cys Leu Leu Tyr				
	385	390	395			400
Asp Leu Ser Thr Thr	Ala Leu Glu His Ile	Leu Leu Phe Ser Leu Gly				
	405	410	415			
Ser Cys Asp Leu Pro	Glu Ser His Leu Ser	Asp Leu Ala Ser Arg Gly				
	420	425	430			
Leu Pro Ala Pro Val	Val Leu Glu Phe Asp	Ser Glu Phe Glu Met Leu				
	435	440	445			
Leu Ala Phe Met Thr	Phe Val Lys Gln Tyr	Gly Pro Glu Phe Val Thr				
	450	455	460			
Gly Tyr Asn Ile Ile	Asn Phe Asp Trp Pro	Phe Val Leu Thr Lys Leu				
	465	470	475			480

Thr Glu Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly  
 485 490 495  
 Arg Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys  
 500 505 510  
 Arg Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly  
 515 520 525  
 Ile Ile Thr Asp Lys Val Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val  
 530 535 540  
 Ala Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp  
 545 550 555 560  
 Ile Pro Ala Tyr Tyr Ala Ser Gly Pro Ala Gln Arg Gly Val Ile Gly  
 565 570 575  
 Glu Tyr Cys Val Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys  
 580 585 590  
 Phe Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile  
 595 600 605  
 Asn Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr  
 610 615 620  
 Cys Leu Leu Arg Leu Ala Gly Gln Lys Gly Phe Ile Leu Pro Asp Thr  
 625 630 635 640  
 Gln Gly Arg Phe Arg Gly Leu Asp Lys Glu Ala Pro Lys Arg Pro Ala  
 645 650 655  
 Val Pro Arg Gly Glu Gly Glu Arg Pro Gly Asp Gly Asn Gly Asp Glu  
 660 665 670  
 Asp Lys Asp Asp Asp Glu Asp Gly Asp Glu Asp Gly Asp Glu Arg Glu  
 675 680 685  
 Glu Val Ala Arg Glu Thr Gly Gly Arg His Val Gly Tyr Gln Gly Ala  
 690 695 700  
 Arg Val Leu Asp Pro Thr Ser Gly Phe His Val Asp Pro Val Val Val  
 705 710 715 720  
 Phe Asp Phe Ala Ser Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu  
 725 730 735  
 Cys Phe Ser Thr Leu Ser Leu Arg Pro Glu Ala Val Ala His Leu Glu  
 740 745 750  
 Ala Asp Arg Asp Tyr Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe  
 755 760 765  
 Phe Val Lys Ala His Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg  
 770 775 780  
 Asp Trp Leu Ala Met Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser  
 785 790 795 800  
 Pro Pro Glu Glu Ala Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys  
 805 810 815

Val Val Cys Asn Ser Val Tyr Gly Phe Thr Gly Ala Gln His Gly Leu  
                     820                    825                    830  
 Leu Pro Cys Leu His Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu  
                     835                    840                    845  
 Met Leu Leu Ala Thr Arg Ala Tyr Val His Ala Arg Trp Ala Glu Phe  
                     850                    855                    860  
 Asp Gln Leu Leu Ala Asp Phe Pro Glu Ala Ala Gly Met Arg Ala Pro  
                     865                    870                    875                    880  
 Gly Pro Tyr Ser Met Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe  
                     885                    890                    895  
 Val Leu Cys Arg Gly Leu Thr Ala Ala Gly Leu Val Ala Met Gly Asp  
                     900                    905                    910  
 Lys Met Ala Ser His Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys  
                     915                    920                    925  
 Leu Glu Cys Glu Lys Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys  
                     930                    935                    940  
 Lys Tyr Ile Gly Val Ile Cys Gly Gly Lys Met Leu Ile Lys Gly Val  
                     945                    950                    955                    960  
 Asp Leu Val Arg Lys Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg  
                     965                    970                    975  
 Ala Leu Val Asp Leu Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala  
                     980                    985                    990  
 Ala Ala Leu Ala Glu Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu  
                     995                    1000                    1005  
 Pro Glu Gly Leu Gln Ala Phe Gly Ala Val Leu Val Asp Ala His  
                     1010                    1015                    1020  
 Arg Arg Ile Thr Asp Pro Glu Arg Asp Ile Gln Asp Phe Val Leu  
                     1025                    1030                    1035  
 Thr Ala Glu Leu Ser Arg His Pro Arg Ala Tyr Thr Asn Lys Arg  
                     1040                    1045                    1050  
 Leu Ala His Leu Thr Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala  
                     1055                    1060                    1065  
 Gln Val Pro Ser Ile Lys Asp Arg Ile Pro Tyr Val Ile Val Ala  
                     1070                    1075                    1080  
 Gln Thr Arg Glu Val Glu Glu Thr Val Ala Arg Leu Ala Ala Leu  
                     1085                    1090                    1095  
 Arg Glu Leu Asp Ala Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro  
                     1100                    1105                    1110  
 Ala Ala Leu Pro Ser Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser  
                     1115                    1120                    1125  
 His Ala Asp Pro Pro Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu

1130	1135	1140
Val Ser Glu Leu Ala Glu Asp	Pro Gly Tyr Ala Ile	Ala Arg Gly
1145	1150	1155
Val Pro Leu Asn Thr Asp Tyr	Tyr Phe Ser His Leu	Leu Gly Ala
1160	1165	1170
Ala Cys Val Thr Phe Lys Ala	Leu Phe Gly Asn Asn	Ala Lys Ile
1175	1180	1185
Thr Glu Ser Leu Leu Lys Arg	Phe Ile Pro Glu Thr	Trp His Pro
1190	1195	1200
Pro Asp Asp Val Ala Ala Arg	Leu Arg Ala Ala Gly	Phe Gly Pro
1205	1210	1215
Ala Gly Ala Gly Ala Thr Ala	Glu Glu Thr Arg Arg	Met Leu His
1220	1225	1230
Arg Ala Phe Asp Thr Leu Ala		
1235	1240	

&lt;210&gt; 5

&lt;211&gt; 3708

&lt;212&gt; DNA

&lt;213&gt; herpes simplex

&lt;400&gt; 5

```

atgtttttccg gtggcgggcgg cccgctgtcc cccggaggaa agtcggcgcc cagggcgggc      60
tccgggtttt ttgcgcccgc cgccctcgc ggagccggcc ggggaccccc gccttgtttg      120
aggcaaaact ttacaaccc ctacctcgcc ccagtcggga cgcaacagaa gccgaccggg      180
ccaaccacgc gccatacgta ctatagcgaa tgcgatgaat ttcgattcat cgcgcccgcg      240
gtgctggacg aggatgcccc cccggagaag cgcgcggggg tgcacgacgg tcacctcaag      300
cgcgccccca aggtgtactg cggggggggac gagcgcgacg tcctccgcgt cgggtcgggc      360
ggcttctggc cgcgcgctc gcgcctgtgg ggcggcgtgg accacgcccc ggcggggttc      420
aaccaccacg tcaccgtctt tcacgtgtac gacatcctgg agaacgtgga gcacgcgtac      480
ggcatgcgcg cggcccagtt ccacgcgcgg tttatggacg ccatcacacc gacggggacc      540
gtcatcacgc tcctgggcct gactccggaa ggccaccggg tggccgttca cgtttacggc      600
acgcggcagt acttttacat gaacaaggag gaggttgaca ggcacctaca atgccgcgcc      660
ccacgagatc tetgcgagcg catggccgcg gccctgcgcg agtccccggg cgcgtcgttc      720
cgcgccatct ccgcggacca cttcgaggcg gaggtggtgg agcgcaccga cgtgtactac      780
tacgagacgc gccccgctct gttttaccgc gtctacgtcc gaagcgggcg cgtgctgtcg      840
tacctgtgcg acaacttctg cccggccatc aagaagtacg aggggtgggt cgacgccacc      900
acccggttca tcctggacaa ccccggttc gtcaccttcg gctggtaccg tctcaaaccg      960
ggccggaaca acacgctagc ccagccgcgg gccccgatgg ccttcgggac atccagcgac     1020

```

gtcagagttta actgtacggc ggacaacctg gccatcgagg ggggcatgag cgacctaccg	1080
gcatacaagc tcatgtgctt cgatatcgaa tgcaaggcgg ggggggagga cgagctggcc	1140
tttccggtgg cggggcaccc ggaggacctg gttattcaga tatcctgtct gctctacgac	1200
ctgtccacca ccgccctgga gcacgtcctc ctgttttcgc tcggttcctg cgacctcccc	1260
gaatcccacc tgaacgagct ggcgccagg ggctgccca cggccgtggg tctggaattc	1320
gacagcgaat tcgagatgct gttggccttc atgacccttg tgaaacagta cggccccgag	1380
ttcgtgaccg ggtacaacat catcaacttc gactggccct tcttgctggc caagttgacg	1440
gacatttaca aggtccccct ggacgggtac ggccgcatga acggccgggg cgtgtttcgc	1500
gtgtgggaca taggccagag ccacttcag aagcgcagca agataaagg gaacggcatg	1560
gtgaacatcg acatgtacgg gatcataacc gacaagatca agctctcgag ctacaagctc	1620
aacgccgtgg ccgaagccgt cctgaaggac aagaagaagg acctgagcta tcgcgacatc	1680
ccgcctact acgccgccgg gcccgcgcaa cgcgggggta tcggcgagta ctgcatacag	1740
gattccctgc tgggtgggcca gctgtttttt aagtttttgc cccatctgga gctctcggcc	1800
gtcgcgcgct tggcggttat taacatcacc cgcaccatct acgacggcca gcagatccgc	1860
gtctttacgt gcctgctgcg cctggccgac cagaagggtt ttattctgcc ggacaccag	1920
gggcgattta gggcgccgg gggggaggcg cccaagcgtc cggccgcagc ccgggaggac	1980
gaggagcggc cagaggagga gggggaggac gaggacgaac gcgaggaggg cgggggcgag	2040
cgggagccgg agggcgcgcg ggagaccgcc ggccggcacg tggggtacca gggggccagg	2100
gtccttgacc ccacttcgg gtttcacgtg aaccccgtag tgggtgtcga ctttgccagc	2160
ctgtacccca gcatcatcca ggcccacaac ctgtgcttca gcacgtctc cctgagggcc	2220
gacgcagtgg cgcacctgga ggcgggcaag gactacctg agatcgaggg gggggggcga	2280
cggctgttct tcgtcaaggc tcacgtgcga gagagcctcc tcagcatcct cctgcgggac	2340
tggctcgcca tgcgaaagca gatccgctcg cggattcccc agagcagccc cgaggaggcc	2400
gtgctcctgg acaagcagca ggccgccatc aaggctcgtg gtaactcggg gtacgggttc	2460
acgggagcgc agcacggact cctgccgtgc ctgcacgttg ccgcgacggg gacgaccatc	2520
ggccgcgaga tgctgctgc gacccgcgag tacgtccacg cgcgctgggc ggccttcgaa	2580
cagctcctgg ccgatttccc ggaggcgcc gacatgcgcg cccccgggcc ctattccatg	2640
cgcacatct acggggacac ggactccata tttgtgctgt gccggggcct cagggccgcc	2700
gggctgacgg ccatgggcga caagatggcg agccacatct cgcgcgcgct gtttctgccc	2760
cccatcaaac tcgagtgcga aaagacgttc accaagctgc tgctgatcgc caagaaaaag	2820
tacatcggcg tcactacgg gggtaagatg ctcatcaagg gcgtggatct ggtgcgcaaa	2880

```

aacaactgcg cgtttatcaa ccgcacctcc agggccctgg tcgacctgct gttttacgac 2940
gataccgtat ccggagcggc cgccgcgtta gccgagcgcc ccgcagagga gtggctggcg 3000
cgacccctgc ccgagggact gcaggcggtc gggggcgtcc tcgtagacgc ccatcggcgc 3060
atcaccgacc cggagagggga catccaggac tttgtcctca ccgccgaact gagcagacac 3120
ccgcgcgcgt acaccaacaa ggcctggcc cacctgacgg tgtattacaa gtcctatggcc 3180
cgccgcgcgc aggtcccgtc catcaaggac cggatcccg acgtgatcgt ggcccagacc 3240
cgcgaggtag aggagacggt cgcgcggtc gccgccctcc gcgagctaga cgccgccgcc 3300
ccaggggacg agcccgcgcc ccccgcggcc ctgccctccc cggccaagcg ccccgggag 3360
acgccgtcgc atgccgaccc cccgggaggc gcgtccaagc cccgcaagct gctggtgtcc 3420
gagctggccg aggatccgc atacgccatt gccacggcg tcgccctgaa cacggactat 3480
tactttctccc acctgttggg ggcggcgtgc gtgacattca aggccctggt tgggaataac 3540
gccaagatca ccgagagtct gttaaaaagg ttattcccg aagtgtggca cccccggac 3600
gacgtggccg cgcggtccg ggccgcaggg ttccgggcgg tgggtgccgg cgctacggcg 3660
gaggaaactc gtcgaatggt gcatagagcc ttgtatactc tagcatga 3708

```

```

<210> 6
<211> 1235
<212> PRT
<213> herpes simplex

```

```

<400> 6

```

```

Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala
1          5          10          15
Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala
20          25          30
Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr
35          40          45
Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg
50          55          60
His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg
65          70          75          80
Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp
85          90          95
Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg
100         105         110
Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg
115         120         125
Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val

```

130	135	140
Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr 145	150	155
Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr 165	170	175
Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His 180	185	190
Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn 195	200	205
Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu 210	215	220
Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe 225	230	235
Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr 245	250	255
Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr 260	265	270
Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro 275	280	285
Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile 290	295	300
Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro 305	310	315
Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly 325	330	335
Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile 340	345	350
Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp 355	360	365
Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala 370	375	380
Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp 385	390	395
Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser 405	410	415
Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu 420	425	430
Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu 435	440	445
Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly 450	455	460

Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495  
 Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560  
 Pro Ala Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765  
 Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780  
 Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800

Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Ala Gln His Gly Leu Leu Pro Cys Leu His  
 820 825 830  
 Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr  
 835 840 845  
 Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala  
 850 855 860  
 Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly  
 885 890 895  
 Leu Thr Ala Ala Gly Leu Thr Ala Met Gly Asp Lys Met Ala Ser His  
 900 905 910  
 Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys  
 915 920 925  
 Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val  
 930 935 940  
 Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys  
 945 950 955 960  
 Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu  
 965 970 975  
 Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu  
 980 985 990  
 Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln  
 995 1000 1005  
 Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp  
 1010 1015 1020  
 Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser  
 1025 1030 1035  
 Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr  
 1040 1045 1050  
 Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile  
 1055 1060 1065  
 Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val  
 1070 1075 1080  
 Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala  
 1085 1090 1095  
 Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser  
 1100 1105 1110  
 Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser His Ala Asp Pro Pro

1115	1120	1125
Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala		
1130	1135	1140
Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr		
1145	1150	1155
Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe		
1160	1165	1170
Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu		
1175	1180	1185
Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Ala		
1190	1195	1200
Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala Gly Ala		
1205	1210	1215
Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr		
1220	1225	1230
Leu Ala		
1235		

<210> 7  
 <211> 3708  
 <212> DNA  
 <213> herpes simplex

<400> 7  
 atgttttccg gtggcggcgg cccgctgtcc cccggaggaa agtcggcggc cagggcggcg 60  
 tccgggtttt ttgcgcccgc cgccctcgc ggagcggcc ggggacccc gccttgcttg 120  
 aggcaaaact tttacaaccc ctacctcgcc ccagtcggga cgcaacagaa gccgaccggg 180  
 ccaaccacgc gccatacgta ctatagcgaa tgcgatgaat ttcgattcat cgcgccgagg 240  
 gtgctggacg aggatgcccc cccggagaag cgcgccgggg tgcacgacgg tcacctcaag 300  
 cgcgccccca aggtgtactg cggggggggac gagcgcgacg tcctccgcgt cgggtcgggc 360  
 ggcttctggc cgcggcgctc gcgcctgtgg ggcggcggtg accacgccc ggcgggggtc 420  
 aacccacccg tcaccgtctt tcacgtgtac gacatcctgg agaacgtgga gcacgcgtac 480  
 ggcattgcgc cgccccagtt ccacgcgcgg tttatggacg ccatcacacc gacggggacc 540  
 gtcattcacgc tcctgggcct gactccggaa ggccaccggg tggccgttca cgtttacggc 600  
 acgcggcagt acttttacat gaacaaggag gaggtcgaca ggcacctaca atgccgcgcc 660  
 ccacgagatc tctgcgagcg catggccgcg gccctgcgcg agtccccggg cgcgtcgttc 720  
 cgcggcattt ccgcggacca cttcgaggcg gaggtggtgg agcgcaccga cgtgtactac 780  
 tacgagacgc gcccgcctct gttttaccgc gtctacgtcc gaagcggggc cgtgctgtcg 840  
 tacctgtgcg acaacttctg cccggccatc aagaagtacg aggggtggggc cgacgccacc 900

acccggttca tcctggacaa ccccggttc gtcaccttcg gctggtaccg tctcaaaccg 960  
 ggccggaaca acacgctagc ccagccgcgg gcccgatgg ccttcgggac atccagcgac 1020  
 gtcgagttta actgtacggc ggacaacctg gccatcgagg ggggcatgag cgacctaccg 1080  
 gcatacaagc tcatgtgctt cgatatcgaa tgcaaggcgg ggggggagga cgagctggcc 1140  
 tttccggtgg cggggcaccg ggaggacctg gtcattccaga taccctgtct gctctacgac 1200  
 ctgtccacca ccgccctgga gcacgtcctc ctgttttcgc tcggttcctg cgacctcccc 1260  
 gaatccacc tgaacgagct ggcgccagg ggcctgcca cggcgtggt tctggaattc 1320  
 gacagcgaat tcgagatgct gttggccttc atgaccttg tgaacagta cggccccgag 1380  
 ttcgtgaccg ggtacaacat catcaacttc gactggccct tcttgctggc caagctgacg 1440  
 gacatttaca aggtccccct ggacgggtac ggccgatga acggccgggg cgtgtttcgc 1500  
 gtgtgggaca taggccagag ccacttcag aagcgagca agataaagg gaacggcatg 1560  
 gtgaacatcg acatgtacgg gattataacc gacaagatca agctctcgag ctacaagctc 1620  
 aacgccgtgg ccgaagccgt cctgaaggac aagaagaagg acctgagcta tcgcgacatc 1680  
 cccgcctact acgccgcgg gccgcgcaa cgcggggtga tcggcgagta ctgcatacag 1740  
 gattccctgc tgggtgggcca gctgtttttt aagtttttcg cccatctgga gctctcgcc 1800  
 gtcgcgcgct tggcggttat taacatcacc cgcacctct acgacggcca gcagatccgc 1860  
 gtctttacgt gcctgctgcg cctggccgac cagaagggt ttattctgcc ggacaccag 1920  
 gggcgattta gggcgggcg gggggaggcg cccaagcgtc cggccgcagc ccgggaggac 1980  
 gaggagcggc cagaggagga gggggaggac gaggacgaac gcgaggaggg cgggggcgag 2040  
 cgggagccgg agggcgcgcg ggagaccgcc ggccggcacg tggggtacca gggggccagg 2100  
 gtccttgacc ccacttcgg gtttcatgtg aaccccgtag tgggttcga ctttgccagc 2160  
 ctgtacccca gcatcatcca ggcccacaac ctgtgcttca gcacgtctc cctgagggcc 2220  
 gacgcagtgg cgcacctgga ggccggcaag gactacctg agatcgagg gggggggcga 2280  
 cggctgttct tcgtcaaggc tcacgtgcga gagagcctc tcagatcct cctgcgggac 2340  
 tggctcgcca tgcgaaagca gatccgctcg cggattcccc agagcagccc cgaggaggcc 2400  
 gtgctcctgg acaagcagca ggccgccatc aaggctcgtg gtaactcgtt ttacgggttc 2460  
 acgggagcgc agcacggact cctgccgtgc ctgcacgttg ccgcgacggt gacgaccatc 2520  
 ggccgcgaga tgctgctcgc gaccgcgag tacgtccacg cgcgctgggc ggccttcgaa 2580  
 cagctcctgg ccgatttccc ggaggcggcc gacatgcgcg cccccgggc ctattccatg 2640  
 cgcacatct acggggacac ggactccatc tttgtgtgt gcccggcct caccggccgc 2700  
 gggctgacgg ccgtgggcga caagatggcg agccacatct cgcgcgcgct gtttctgtcc 2760

```

cccatcaaac tcgagtgcga aaagacgttc accaagctgc tgctgatcgc caagaaaaag 2820
tacatcggcg tcacttacgg gggtaagatg ctcacaaagg gcgtggatct ggtgcgcaaa 2880
aacaactgcg cgtttatcaa ccgcacctcc agggccctgg tcgacctgct gttttacgac 2940
gataccgtat ccggagcggc cgccgcgtta gccgagcgcc ccgcagagga gtggctggcg 3000
cgaccctgc ccgagggact gcaggcggtc ggggcccgtc tcgtagacgc ccatcggcg 3060
atcacccgacc cggagagggga catccaggac tttgtcctca ccgccgaact gagcagacac 3120
ccgcgcgcgt acaccaacaa gcgcctggcc cacctgacgg tgtattacaa gctcatggcc 3180
cgccgcgcgc aggtcccgtc catcaaggac cggatcccgt acgtgatcgt ggcccagacc 3240
cgcgaggtag aggagacggc cgcgcggtg gccgccctcc gcgagctcga cgccgccgcc 3300
ccaggggacg agcccgcgcc ccccgcggcc ctgccctccc cggccaagcg ccccgggag 3360
acgccgttgc atgccgacc cccgggaggg gcgtccaagc cccgcaagct gctggtgtcc 3420
gagctggccg aggatccgc atacgccatt gccacggcg tcgccctgaa cacggactat 3480
tacttctccc acctgttggg gcgcgcgtgc gtgacattca aggccctgtt tgggaataac 3540
gccaagatca ccgagagtct gttaaaaagg tttattcccg aagtgtggca ccccccggac 3600
gacgtggccg cgcggtccg ggccgcaggg ttccggggcg tgggtgccgg cgctacggcg 3660
gaggaaactc gtcgaatgtt gcatagagcc tttgatactc tagcatga 3708

```

```

<210> 8
<211> 1235
<212> PRT
<213> herpes simplex

```

```

<400> 8

```

```

Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala
1           5           10          15

Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala
20          25          30

Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr
35          40          45

Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg
50          55          60

His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg
65          70          75          80

Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp
85          90          95

Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg
100         105         110

Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg

```

115	120	125
Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val 130	135	140
Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr 145	150	155
Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr 165	170	175
Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His 180	185	190
Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn 195	200	205
Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu 210	215	220
Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe 225	230	235
Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr 245	250	255
Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr 260	265	270
Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro 275	280	285
Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile 290	295	300
Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro 305	310	315
Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly 325	330	335
Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile 340	345	350
Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp 355	360	365
Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala 370	375	380
Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp 385	390	395
Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser 405	410	415
Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu 420	425	430
Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu 435	440	445

Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly  
 450 455 460  
 Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495  
 Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560  
 Pro Ala Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Gly Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765  
 Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780

Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800  
 Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Ala Gln His Gly Leu Leu Pro Cys Leu His  
 820 825 830  
 Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr  
 835 840 845  
 Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala  
 850 855 860  
 Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly  
 885 890 895  
 Leu Thr Ala Ala Gly Leu Thr Ala Val Gly Asp Lys Met Ala Ser His  
 900 905 910  
 Ile Ser Arg Ala Leu Phe Leu Ser Pro Ile Lys Leu Glu Cys Glu Lys  
 915 920 925  
 Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val  
 930 935 940  
 Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys  
 945 950 955 960  
 Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu  
 965 970 975  
 Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu  
 980 985 990  
 Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln  
 995 1000 1005  
 Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp  
 1010 1015 1020  
 Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser  
 1025 1030 1035  
 Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr  
 1040 1045 1050  
 Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile  
 1055 1060 1065  
 Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val  
 1070 1075 1080  
 Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala  
 1085 1090 1095  
 Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser

1100	1105	1110
Pro Ala Lys Arg Pro Arg Glu Thr Pro Leu His Ala Asp Pro Pro		
1115	1120	1125
Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala		
1130	1135	1140
Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr		
1145	1150	1155
Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe		
1160	1165	1170
Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu		
1175	1180	1185
Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Ala		
1190	1195	1200
Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala Gly Ala		
1205	1210	1215
Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr		
1220	1225	1230
Leu Ala		
1235		

<210> 9  
 <211> 3708  
 <212> DNA  
 <213> herpes simplex

<400> 9  
 atgtttttccg gtggcgggcgg cccgctgtcc cccggaggaa agtcggcggc cagggcgggc 60  
 tccgggtttt ttgcgcccgc cggccctcgc ggagccggcc ggggaccccc gccttgtttg 120  
 aggcaaaact tttacaaccc ctacctcgcc ccagtcggga cgcaacagaa gccgaccggg 180  
 ccaaccacgc gccatacgta ctatagcgaa tgcgatgaat ttcgattcat cgcccccgcg 240  
 gtgctggacg aggatgcccc cccggagaag cgcgcggggg tgcacgacgg tcacctcaag 300  
 cgcgccccca aggtgtactg cggggggggac gagcgcgacg tcctccgcgt cgggtcgggc 360  
 ggcttctggc cgcggcgctc gcgcctgtgg ggcggcggtg accacgcccc ggcgggggttc 420  
 aacccacccg tcacgtcttt tcacgtgtat gacatcctgg agaacgtgga gcacgcgtac 480  
 ggcacgcgcg cggcccagtt ccacgcgcgg tttatggacg ccatcacacc gacgggggacc 540  
 gtcacacgcg tcttgggcct gactccggaa ggccaccggg tggccgttca cgtttacggc 600  
 acgcggcagc actttttacat gaacaaggag gaggttgaca ggcacctaca atgccgcgcc 660  
 ccacgagatc tctgcgagcg catggccgcg gccctgcgcg agtccccggg cgcgtcgttc 720  
 cgcggcatct ccgcggacca cttcgaggcg gaggtggtgg agcgcaccga cgtgtactac 780  
 tacgagacgc gccccgctct gttttaccgc gtctacgtcc gaagcggggc cgtgctgtcg 840

tacctgtgcg acaacttctg cccggccatc aagaagtacg aggggtggggt cgacgccacc	900
acccggttca tcctggacaa ccccggttc gtcaccttcg gctggtaccg tctcaaaccg	960
ggccggaaca acacgctagc ccagccgcgg gccccgatgg ccttcgggac atccagcgat	1020
gtcgagtta actgtacggc ggacaacctg gccatcgagg ggggcatgag cgacctaccg	1080
gcatacaagc tcatgtgctt cgatatcgaa tgcaaggcgg ggggggagga cgagctggcc	1140
tttccggtgg cggggcaccg ggaggacctg gtcattccaga tatcctgtct gctctacgac	1200
ctgtccacca ccgccctgga gcacgtcctc ctgttttcgc tcggttcctg cgacctcccc	1260
gaatcccacc tgaacgagct ggccggccagg ggcctgccca cggccgtggg tctggaattc	1320
gacagcgaat tcgagatgct gttggccttc atgaccttg tgaacagta cggccccgag	1380
ttcgtgaccg ggtacaacat aatcaacttc gactggccct tcttgctggc caagctgacg	1440
gacatttaca aggtccccct ggacgggtac ggccgcatga acggccgggg cgtgtttcgc	1500
gtgtgggaca taggccagag ccacttccag aagcgcagca agataaagg gaacggcatg	1560
gtgaacatcg acatgtacgg gattataacc gacaagatca agctctcgag ctacaagctc	1620
aacgccgtgg ccgaagccgt cctgaaggac aagaagaagg acctgagcta tcgcgacatc	1680
cccacctact acgccgccgg gcccgcgcaa cgcgggggtga tcggcgagta ctgcatacag	1740
gattccctgc tgggtgggcca gctgtttttt aagtttttgc cccatctgga gctctcgcc	1800
gtcgcgcgct tggcgggtat taacatcacc cgcacctct acgacggcca gcagatccgc	1860
gtctttacgt gcctgctgcg cctggccgac cagaagggt ttattctgcc ggacaccag	1920
gggcgattta ggggcgccgg gggggaggcg cccaagcgtc cggccgcagc ccgggaggac	1980
gaggagcggc cagaggagga gggggaggac gagaacgaac gcgaggaggg cgggggcgag	2040
cgggagccgg agggcgcgcg ggagaccgcc ggccggcacg tggggtacca gggggccagg	2100
gtccttgacc ccacttccgg gtttcacgtg aaccccgtagg tgggtgtcga ctttgccagc	2160
ctgtacccca gcatcatcca ggcccacaac ctgtgcttca gcacgtctc cctgagggcc	2220
gacgcagtgg cgcacctgga ggccggcaag gactacctgg agatcgaggg gggggggcga	2280
cggctgttct tcgtcaaggc tcacgtgcga gagagcctcc tcagcatcct cctgcgggac	2340
tggctcgcca tgcgaaagca gatccgctcg cggattcccc agagcagccc cgaggaggcc	2400
gtgctcctgg acaagcagca ggccgccatc aaggctcgtg gtaactcggg ttacgggttc	2460
acgggagcgc agcacggact cctgccgtgc ctgcacgttg ccgcgacggg gacgaccatc	2520
ggccgcgaga tgctgctcgc gaccgcgag tacgtccacg cgcgctgggc ggccttcgaa	2580
cagctcctgg ccgatttccc ggaggcgggc gacatgcgcg cccccgggcc ctattccatg	2640
cgcacatct acggggacac ggactccata tttgtgctgt gcccgccct caccggccgc	2700

```

gggctgacgg ccgtagggcga caagatggcg agccacatct cgcgcgcgct gtttctgccc 2760
cccatcaaac tcgagtgcga aaagacgttc accaagctgc tgctgatcgc caagaaaaag 2820
tacatcggcg tcactctacgg gggtaagatg ctcatcaagg gcgtggatct ggtgcgcaaa 2880
aacaactgcg cgtttatcaa cgcacacctc agggccctgg tcgacctgct gttttacgac 2940
gataccgtat cggagagcgg cgcgcgttta gccgagcgcc cgcagagga gtggctggcg 3000
cgacctctgc ccgagggact gcaggcggtc ggggccgtcc tcgtagacgc ccatcggcgc 3060
atcaccgacc cggagagggga catccaggac tttgttctca cgcgccgaact gagcagacac 3120
cgcgcgcgct acaccaacaa ggcctggcc cacctgacgg tgtattacaa gctcatggcc 3180
cgcgcgcgcg aggtcccgtc catcaaggac cggatcccgt acgtgatcgt ggcccagacc 3240
cgcgaggtag aggagacggt cgcgcggctg gccgccctcc gcgagctaga cgcgcgcgcg 3300
ccaggggaag agcccgcgcc ccccgcggcc ctgccctccc cggccaagcg ccccggggag 3360
acgccgtcgc ctgccgaccc cccgggaggg gcgtccaagc cccgcaagct gctggtgtcc 3420
gagctggccg aggatcccgc atacgccatt gccacggcg tcgccctgaa cacggactat 3480
tacttctccc acctgttggg ggcggcgtgc gtgacattca aggccctgtt tgggaataac 3540
gccaagatca ccgagagtct gttaaaaagg ttattcccg aagtgtggca cccccggac 3600
gacgtggccg cgcggctccg gaccgcaggg ttcggggcgg tgggtgccgg cgctacggcg 3660
gaggaaactc gtcgaatggt gcatagagcc tttgatactc tagcatga 3708

```

<210> 10  
 <211> 1235  
 <212> PRT  
 <213> herpes simplex

<400> 10

```

Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala
1             5             10             15

Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala
20             25             30

Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr
35             40             45

Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg
50             55             60

His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg
65             70             75             80

Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp
85             90             95

Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg

```

100					105					110						
Asp	Val	Leu	Arg	Val	Gly	Ser	Gly	Gly	Phe	Trp	Pro	Arg	Arg	Ser	Arg	
115					120					125						
Leu	Trp	Gly	Gly	Val	Asp	His	Ala	Pro	Ala	Gly	Phe	Asn	Pro	Thr	Val	
130					135					140						
Thr	Val	Phe	His	Val	Tyr	Asp	Ile	Leu	Glu	Asn	Val	Glu	His	Ala	Tyr	
145					150					155					160	
Gly	Met	Arg	Ala	Ala	Gln	Phe	His	Ala	Arg	Phe	Met	Asp	Ala	Ile	Thr	
165					170					175						
Pro	Thr	Gly	Thr	Val	Ile	Thr	Leu	Leu	Gly	Leu	Thr	Pro	Glu	Gly	His	
180					185					190						
Arg	Val	Ala	Val	His	Val	Tyr	Gly	Thr	Arg	Gln	Tyr	Phe	Tyr	Met	Asn	
195					200					205						
Lys	Glu	Glu	Val	Asp	Arg	His	Leu	Gln	Cys	Arg	Ala	Pro	Arg	Asp	Leu	
210					215					220						
Cys	Glu	Arg	Met	Ala	Ala	Ala	Leu	Arg	Glu	Ser	Pro	Gly	Ala	Ser	Phe	
225					230					235					240	
Arg	Gly	Ile	Ser	Ala	Asp	His	Phe	Glu	Ala	Glu	Val	Val	Glu	Arg	Thr	
245					250					255						
Asp	Val	Tyr	Tyr	Tyr	Glu	Thr	Arg	Pro	Ala	Leu	Phe	Tyr	Arg	Val	Tyr	
260					265					270						
Val	Arg	Ser	Gly	Arg	Val	Leu	Ser	Tyr	Leu	Cys	Asp	Asn	Phe	Cys	Pro	
275					280					285						
Ala	Ile	Lys	Lys	Tyr	Glu	Gly	Gly	Val	Asp	Ala	Thr	Thr	Arg	Phe	Ile	
290					295					300						
Leu	Asp	Asn	Pro	Gly	Phe	Val	Thr	Phe	Gly	Trp	Tyr	Arg	Leu	Lys	Pro	
305					310					315					320	
Gly	Arg	Asn	Asn	Thr	Leu	Ala	Gln	Pro	Arg	Ala	Pro	Met	Ala	Phe	Gly	
325					330					335						
Thr	Ser	Ser	Asp	Val	Glu	Phe	Asn	Cys	Thr	Ala	Asp	Asn	Leu	Ala	Ile	
340					345					350						
Glu	Gly	Gly	Met	Ser	Asp	Leu	Pro	Ala	Tyr	Lys	Leu	Met	Cys	Phe	Asp	
355					360					365						
Ile	Glu	Cys	Lys	Ala	Gly	Gly	Glu	Asp	Glu	Leu	Ala	Phe	Pro	Val	Ala	
370					375					380						
Gly	His	Pro	Glu	Asp	Leu	Val	Ile	Gln	Ile	Ser	Cys	Leu	Leu	Tyr	Asp	
385					390					395					400	
Leu	Ser	Thr	Thr	Ala	Leu	Glu	His	Val	Leu	Leu	Phe	Ser	Leu	Gly	Ser	
405					410					415						
Cys	Asp	Leu	Pro	Glu	Ser	His	Leu	Asn	Glu	Leu	Ala	Ala	Arg	Gly	Leu	
420					425					430						

Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu  
 435 440 445  
 Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly  
 450 455 460  
 Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495  
 Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560  
 Pro Thr Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asn  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Thr Ser Gly Phe His Val Asn Pro Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765

Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780  
 Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800  
 Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Ala Gln His Gly Leu Leu Pro Cys Leu His  
 820 825 830  
 Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr  
 835 840 845  
 Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala  
 850 855 860  
 Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly  
 885 890 895  
 Leu Thr Ala Ala Gly Leu Thr Ala Val Gly Asp Lys Met Ala Ser His  
 900 905 910  
 Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys  
 915 920 925  
 Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val  
 930 935 940  
 Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys  
 945 950 955 960  
 Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu  
 965 970 975  
 Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu  
 980 985 990  
 Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln  
 995 1000 1005  
 Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp  
 1010 1015 1020  
 Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser  
 1025 1030 1035  
 Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr  
 1040 1045 1050  
 Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile  
 1055 1060 1065  
 Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val  
 1070 1075 1080  
 Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala

1085	1090	1095
Ala Ala Pro Gly Asp Glu Pro	Ala Pro Pro Ala Ala	Leu Pro Ser
1100	1105	1110
Pro Ala Lys Arg Pro Arg Glu	Thr Pro Ser Pro Ala	Asp Pro Pro
1115	1120	1125
Gly Gly Ala Ser Lys Pro Arg	Lys Leu Leu Val Ser	Glu Leu Ala
1130	1135	1140
Glu Asp Pro Ala Tyr Ala Ile	Ala His Gly Val Ala	Leu Asn Thr
1145	1150	1155
Asp Tyr Tyr Phe Ser His Leu	Leu Gly Ala Ala Cys	Val Thr Phe
1160	1165	1170
Lys Ala Leu Phe Gly Asn Asn	Ala Lys Ile Thr Glu	Ser Leu Leu
1175	1180	1185
Lys Arg Phe Ile Pro Glu Val	Trp His Pro Pro Asp	Asp Val Ala
1190	1195	1200
Ala Arg Leu Arg Thr Ala Gly	Phe Gly Ala Val Gly	Ala Gly Ala
1205	1210	1215
Thr Ala Glu Glu Thr Arg Arg	Met Leu His Arg Ala	Phe Asp Thr
1220	1225	1230
Leu Ala		
1235		

<210> 11  
 <211> 3729  
 <212> DNA  
 <213> herpes simplex

<400> 11  
 atgtttttca acccgatatct gagcgggcggc gtgaccggcg gtgcgggtcgc ggggtggccgg 60  
 cgtcagcgtt cgcagcccgg ctccgcgcag ggctcgggca agcggccgcc acagaaacag 120  
 tttttgcaga tcgtgccgcg aggtgtcatg ttcgacggtc agacgggggtt gatcaagcat 180  
 aagacgggac ggctgcctct catgttctat cgagagatta aacatttggt gagtcatgac 240  
 atggtttggc cgtgtccttg gcgcgagacc ctgggtgggtc gcgtgggtggg acctattcgt 300  
 tttcacacct acgatcagac ggacgccgtg ctcttcttcg actcgcccga aaacgtgtcg 360  
 ccgcgctatc gtcagcatct ggtgccttcg gggaacgtgt tgcgtttctt cggggccaca 420  
 gaacacgggt acagtatctg cgtcaacgtt ttcgggcagc gcagctactt ttactgtgag 480  
 tacagcgaca ccgataggct gcgtgaggtc attgccagcg tgggcgaact agtgcccga 540  
 ccgcggacgc catacggcgt gtctgtcacg ccggccacca agacctccat ctatgggtac 600  
 gggacgcgac ccgtgcccga tttgcagtgt gtgtctatca gcaactggac catggccaga 660  
 aaaatcggcg agtatctgct ggagcagggt tttcccggtg acgaggtccg tgtggatccg 720

ctgacgcgtt tggatcatcga tgggaggatc accacgttcg gctgggtgctc cgtgaatcgt	780
tacgactggc ggcagcaggg tcgcgctcgc acttgatgata tcgaggtaga ctgcgatgac	840
tctgacctgg tggctgtgcc cgacgacagc tcgtggccgc gctatcgatg cctgtccttc	900
gatatcgagt gcatgagcgg cgagggtggg tttccctgcg ccgagaagtc cgatgacatt	960
gtcattcaga tctcgtgcgt gtgctacgag acggggggaa acaccgccgt ggatcagggg	1020
atcccaaacg ggaacgatgg tcggggctgc acttcggagg gtgtgatctt tgggcactcg	1080
ggctttcatc tctttacgat cggcacctgc gggcagggtg gccagacgt ggacgtctac	1140
gagttccctt ccgaatacga gctgctgctg ggctttatgc ttttctttca acggtacgcg	1200
ccggcctttg tgaccggtta caacatcaac tcttttgact tgaagtacat cctcacgcgt	1260
ctcgagtacc tgtataaggt ggactcgcag cgcttctgca agttgcctac ggcgcagggc	1320
ggccgtttct ttttacacag ccccgccgtg gggttttaagc ggcagtacgc cgccgctttt	1380
ccctcggtt ctcaacaac tccggccagc acggccgcc ccaagggtgta tattgcggtt	1440
tcggtgggta tcgacatgta ccctgatgac atggccaaga ctaactcgcc caactataag	1500
ctcaacacta tggccgagct ttacctgcgg caacgcaagg atgacctgtc ttacaaggac	1560
atcccgctt gtttcgtggc taatgccgag ggccgcgcc aggtaggccg ttactgtctg	1620
caggacgccg tattggtgcg cgatctgttc aacaccatta attttacta cgaggccggg	1680
gccatcgccg ggctggctaa aattccgttg cggcgtgtca tctttgacgg acagcagatc	1740
cgtatctaca cctcgctgct ggacgagtg gcctgccgg attttatcct gcccaaccac	1800
tacagcaaag gtacgacggt gcccgaacg aatagcgttg ctgtgtcacc taacgctgct	1860
atcatctcta ccgccgctgt gcccggcgac ggggttctg tggcggtat gtttcagatg	1920
tcgccgccct tgcaatctgc gccgtccagt caggacggcg tttcaccgg ctccggcagt	1980
aacagtagta gcagcgtcgg cgttttcagc gtcggctccg gcagtagtgg cggcgtcggc	2040
gtttccaacg acaatcacgg cgccggcgg actgcggcg tttcgtacca gggcgccacg	2100
gtgtttgagc ccgaggtggg ttactacaac gacccgtgg ccgtgttcga ctttgccagc	2160
ctctaccctt ccatcatcat ggcccacaac ctctgtact ccacctgct ggtgccgggt	2220
ggcgagtacc ctgtggacct cgccgacgta tacagcgtca cgctagagaa cggcgtgacc	2280
caccgctttg tgcgtgcttc ggtgcgcgtc tcggtgtctt cggaactgct caacaagtgg	2340
gtttcgagc ggcgtgccgt gcgcgaatgc atgcgcgagt gtcaagacct tgtgcgccgt	2400
atgtgtctcg acaaggaaca gatggcgctc aaagtaacgt gcaacgcttt ctacggtttt	2460
accggcgccg tgaacgggat gatgccgtgt ctgccatcg ccgccagcat cacgcgcac	2520
ggtcgcgaca tgctagagcg caccggcgcg ttcacaaag acaacttttc agagccgtgt	2580

```

tttttgcaca attttttttaa tcaggaagac tatgtagtgg gaacgcggga gggggattcg 2640
gaggagagca ggcggttacc ggaggggctc gaaacatcgt cagggggctc gaacgaacgg 2700
cgggtggagg cgcggtcat ctacggggac acggacagcg tgtttgtccg ctttcgtggc 2760
ctgacgccgc aggctctggt ggcggtggg ccagcctgg cgcactacgt gacggcctgt 2820
ctttttgtgg agcccgtaa gctggagttt gaaaaggctc tcgtctctct tatgatgatc 2880
tgcaagaaac gttacatcgg caaagtggag ggcgctcgg gtctgagcat gaagggcgtg 2940
gatctggtgc gcaagacggc ctgcgagttc gtcaagggcg tcacgcgtga cgtcctctcg 3000
ctgctctttg aggatcgga ggtctcgga gcagccgtgc gcctgtcgg cctctcactc 3060
gatgaagtca agaagtacgg cgtgccacgc ggtttctggc gtatcttacg ccgcttggtg 3120
cagggccgcg acgatctgta cctgcaccgt gtgcgtgtcg aggacctggt gctttcgtcg 3180
gtgctctcta aggacatctc gctgtaccgt caatctaacc tgccgcacat tgccgtcatt 3240
aagcgattgg cggcccgttc tgaggagcta ccctcggtcg gggatcgggt cttttacgtt 3300
ctgacggcgc ccggtgtccg gacggcgccg cagggttctt ccgacaacgg tgattctgta 3360
accgcccggc tggtttcccg gtcggacgcg attgatggca cggacgacga cgctgacggc 3420
ggcggggtag aggagagcaa caggagagga ggagagccgg caaagaagag ggcgcgga 3480
ccaccgtcgg ccgtgtgcaa ctacgaggta gccgaagatc cgagctacgt gcgcgagcac 3540
ggcgtgcca ttcacgccga caagtacttt gagcaggttc tcaaggctgt aactaacgtg 3600
ctgtcgcccg tctttcccg cggcgaaacc gcgcgaagg acaagttttt gcacatgggtg 3660
ctgcccgggc gcttgcactt ggagccggct tttctgccgt acagtgtcaa ggcgcacgaa 3720
tgctgttga 3729

```

<210> 12  
 <211> 1242  
 <212> PRT  
 <213> herpes simplex

<400> 12

```

Met Phe Phe Asn Pro Tyr Leu Ser Gly Gly Val Thr Gly Gly Ala Val
1           5           10           15

Ala Gly Gly Arg Arg Gln Arg Ser Gln Pro Gly Ser Ala Gln Gly Ser
20           25           30

Gly Lys Arg Pro Pro Gln Lys Gln Phe Leu Gln Ile Val Pro Arg Gly
35           40           45

Val Met Phe Asp Gly Gln Thr Gly Leu Ile Lys His Lys Thr Gly Arg
50           55           60

Leu Pro Leu Met Phe Tyr Arg Glu Ile Lys His Leu Leu Ser His Asp
65           70           75           80

```

Met Val Trp Pro Cys Pro Trp Arg Glu Thr Leu Val Gly Arg Val Val  
 85 90 95  
 Gly Pro Ile Arg Phe His Thr Tyr Asp Gln Thr Asp Ala Val Leu Phe  
 100 105 110  
 Phe Asp Ser Pro Glu Asn Val Ser Pro Arg Tyr Arg Gln His Leu Val  
 115 120 125  
 Pro Ser Gly Asn Val Leu Arg Phe Phe Gly Ala Thr Glu His Gly Tyr  
 130 135 140  
 Ser Ile Cys Val Asn Val Phe Gly Gln Arg Ser Tyr Phe Tyr Cys Glu  
 145 150 155 160  
 Tyr Ser Asp Thr Asp Arg Leu Arg Glu Val Ile Ala Ser Val Gly Glu  
 165 170 175  
 Leu Val Pro Glu Pro Arg Thr Pro Tyr Ala Val Ser Val Thr Pro Ala  
 180 185 190  
 Thr Lys Thr Ser Ile Tyr Gly Tyr Gly Thr Arg Pro Val Pro Asp Leu  
 195 200 205  
 Gln Cys Val Ser Ile Ser Asn Trp Thr Met Ala Arg Lys Ile Gly Glu  
 210 215 220  
 Tyr Leu Leu Glu Gln Gly Phe Pro Val Tyr Glu Val Arg Val Asp Pro  
 225 230 235 240  
 Leu Thr Arg Leu Val Ile Asp Arg Arg Ile Thr Thr Phe Gly Trp Cys  
 245 250 255  
 Ser Val Asn Arg Tyr Asp Trp Arg Gln Gln Gly Arg Ala Ser Thr Cys  
 260 265 270  
 Asp Ile Glu Val Asp Cys Asp Val Ser Asp Leu Val Ala Val Pro Asp  
 275 280 285  
 Asp Ser Ser Trp Pro Arg Tyr Arg Cys Leu Ser Phe Asp Ile Glu Cys  
 290 295 300  
 Met Ser Gly Glu Gly Gly Phe Pro Cys Ala Glu Lys Ser Asp Asp Ile  
 305 310 315 320  
 Val Ile Gln Ile Ser Cys Val Cys Tyr Glu Thr Gly Gly Asn Thr Ala  
 325 330 335  
 Val Asp Gln Gly Ile Pro Asn Gly Asn Asp Gly Arg Gly Cys Thr Ser  
 340 345 350  
 Glu Gly Val Ile Phe Gly His Ser Gly Leu His Leu Phe Thr Ile Gly  
 355 360 365  
 Thr Cys Gly Gln Val Gly Pro Asp Val Asp Val Tyr Glu Phe Pro Ser  
 370 375 380  
 Glu Tyr Glu Leu Leu Leu Gly Phe Met Leu Phe Phe Gln Arg Tyr Ala  
 385 390 395 400  
 Pro Ala Phe Val Thr Gly Tyr Asn Ile Asn Ser Phe Asp Leu Lys Tyr

405					410					415					
Ile	Leu	Thr	Arg	Leu	Glu	Tyr	Leu	Tyr	Lys	Val	Asp	Ser	Gln	Arg	Phe
420				425				430							
Cys	Lys	Leu	Pro	Thr	Ala	Gln	Gly	Gly	Arg	Phe	Phe	Leu	His	Ser	Pro
435				440				445							
Ala	Val	Gly	Phe	Lys	Arg	Gln	Tyr	Ala	Ala	Ala	Phe	Pro	Ser	Ala	Ser
450				455				460							
His	Asn	Asn	Pro	Ala	Ser	Thr	Ala	Ala	Thr	Lys	Val	Tyr	Ile	Ala	Gly
465				470				475				480			
Ser	Val	Val	Ile	Asp	Met	Tyr	Pro	Val	Cys	Met	Ala	Lys	Thr	Asn	Ser
				485				490				495			
Pro	Asn	Tyr	Lys	Leu	Asn	Thr	Met	Ala	Glu	Leu	Tyr	Leu	Arg	Gln	Arg
500				505				510							
Lys	Asp	Asp	Leu	Ser	Tyr	Lys	Asp	Ile	Pro	Arg	Cys	Phe	Val	Ala	Asn
515				520				525							
Ala	Glu	Gly	Arg	Ala	Gln	Val	Gly	Arg	Tyr	Cys	Leu	Gln	Asp	Ala	Val
530				535				540							
Leu	Val	Arg	Asp	Leu	Phe	Asn	Thr	Ile	Asn	Phe	His	Tyr	Glu	Ala	Gly
545				550				555				560			
Ala	Ile	Ala	Arg	Leu	Ala	Lys	Ile	Pro	Leu	Arg	Arg	Val	Ile	Phe	Asp
				565				570				575			
Gly	Gln	Gln	Ile	Arg	Ile	Tyr	Thr	Ser	Leu	Leu	Asp	Glu	Cys	Ala	Cys
580				585				590							
Arg	Asp	Phe	Ile	Leu	Pro	Asn	His	Tyr	Ser	Lys	Gly	Thr	Thr	Val	Pro
595				600				605							
Glu	Thr	Asn	Ser	Val	Ala	Val	Ser	Pro	Asn	Ala	Ala	Ile	Ile	Ser	Thr
610				615				620							
Ala	Ala	Val	Pro	Gly	Asp	Ala	Gly	Ser	Val	Ala	Ala	Met	Phe	Gln	Met
625				630				635				640			
Ser	Pro	Pro	Leu	Gln	Ser	Ala	Pro	Ser	Ser	Gln	Asp	Gly	Val	Ser	Pro
				645				650				655			
Gly	Ser	Gly	Ser	Asn	Ser	Ser	Ser	Ser	Val	Gly	Val	Phe	Ser	Val	Gly
660				665				670							
Ser	Gly	Ser	Ser	Gly	Gly	Val	Gly	Val	Ser	Asn	Asp	Asn	His	Gly	Ala
675				680				685							
Gly	Gly	Thr	Ala	Ala	Val	Ser	Tyr	Gln	Gly	Ala	Thr	Val	Phe	Glu	Pro
690				695				700							
Glu	Val	Gly	Tyr	Tyr	Asn	Asp	Pro	Val	Ala	Val	Phe	Asp	Phe	Ala	Ser
705				710				715				720			
Leu	Tyr	Pro	Ser	Ile	Ile	Met	Ala	His	Asn	Leu	Cys	Tyr	Ser	Thr	Leu
				725				730				735			

Leu Val Pro Gly Gly Glu Tyr Pro Val Asp Pro Ala Asp Val Tyr Ser  
 740 745 750  
 Val Thr Leu Glu Asn Gly Val Thr His Arg Phe Val Arg Ala Ser Val  
 755 760 765  
 Arg Val Ser Val Leu Ser Glu Leu Leu Asn Lys Trp Val Ser Gln Arg  
 770 775 780  
 Arg Ala Val Arg Glu Cys Met Arg Glu Cys Gln Asp Pro Val Arg Arg  
 785 790 795 800  
 Met Leu Leu Asp Lys Glu Gln Met Ala Leu Lys Val Thr Cys Asn Ala  
 805 810 815  
 Phe Tyr Gly Phe Thr Gly Ala Leu Asn Gly Met Met Pro Cys Leu Pro  
 820 825 830  
 Ile Ala Ala Ser Ile Thr Arg Ile Gly Arg Asp Met Leu Glu Arg Thr  
 835 840 845  
 Ala Arg Phe Ile Lys Asp Asn Phe Ser Glu Pro Cys Phe Leu His Asn  
 850 855 860  
 Phe Phe Asn Gln Glu Asp Tyr Val Val Gly Thr Arg Glu Gly Asp Ser  
 865 870 875 880  
 Glu Glu Ser Ser Ala Leu Pro Glu Gly Leu Glu Thr Ser Ser Gly Gly  
 885 890 895  
 Ser Asn Glu Arg Arg Val Glu Ala Arg Val Ile Tyr Gly Asp Thr Asp  
 900 905 910  
 Ser Val Phe Val Arg Phe Arg Gly Leu Thr Pro Gln Ala Leu Val Ala  
 915 920 925  
 Arg Gly Pro Ser Leu Ala His Tyr Val Thr Ala Cys Leu Phe Val Glu  
 930 935 940  
 Pro Val Lys Leu Glu Phe Glu Lys Val Phe Val Ser Leu Met Met Ile  
 945 950 955 960  
 Cys Lys Lys Arg Tyr Ile Gly Lys Val Glu Gly Ala Ser Gly Leu Ser  
 965 970 975  
 Met Lys Gly Val Asp Leu Val Arg Lys Thr Ala Cys Glu Phe Val Lys  
 980 985 990  
 Gly Val Thr Arg Asp Val Leu Ser Leu Leu Phe Glu Asp Arg Glu Val  
 995 1000 1005  
 Ser Glu Ala Ala Val Arg Leu Ser Arg Leu Ser Leu Asp Glu Val  
 1010 1015 1020  
 Lys Lys Tyr Gly Val Pro Arg Gly Phe Trp Arg Ile Leu Arg Arg  
 1025 1030 1035  
 Leu Val Gln Ala Arg Asp Asp Leu Tyr Leu His Arg Val Arg Val  
 1040 1045 1050  
 Glu Asp Leu Val Leu Ser Ser Val Leu Ser Lys Asp Ile Ser Leu  
 1055 1060 1065

Tyr Arg Gln Ser Asn Leu Pro His Ile Ala Val Ile Lys Arg Leu  
 1070 1075 1080  
 Ala Ala Arg Ser Glu Glu Leu Pro Ser Val Gly Asp Arg Val Phe  
 1085 1090 1095  
 Tyr Val Leu Thr Ala Pro Gly Val Arg Thr Ala Pro Gln Gly Ser  
 1100 1105 1110  
 Ser Asp Asn Gly Asp Ser Val Thr Ala Gly Val Val Ser Arg Ser  
 1115 1120 1125  
 Asp Ala Ile Asp Gly Thr Asp Asp Asp Ala Asp Gly Gly Gly Val  
 1130 1135 1140  
 Glu Glu Ser Asn Arg Arg Gly Gly Glu Pro Ala Lys Lys Arg Ala  
 1145 1150 1155  
 Arg Lys Pro Pro Ser Ala Val Cys Asn Tyr Glu Val Ala Glu Asp  
 1160 1165 1170  
 Pro Ser Tyr Val Arg Glu His Gly Val Pro Ile His Ala Asp Lys  
 1175 1180 1185  
 Tyr Phe Glu Gln Val Leu Lys Ala Val Thr Asn Val Leu Ser Pro  
 1190 1195 1200  
 Val Phe Pro Gly Gly Glu Thr Ala Arg Lys Asp Lys Phe Leu His  
 1205 1210 1215  
 Met Val Leu Pro Arg Arg Leu His Leu Glu Pro Ala Phe Leu Pro  
 1220 1225 1230  
 Tyr Ser Val Lys Ala His Glu Cys Cys  
 1235 1240

<210> 13  
 <211> 1242  
 <212> PRT  
 <213> herpes simplex

<400> 13

Met Phe Phe Asn Pro Tyr Leu Ser Gly Gly Val Thr Gly Gly Ala Val  
 1 5 10 15  
 Ala Gly Gly Arg Arg Gln Arg Ser Gln Pro Gly Ser Ala Gln Gly Ser  
 20 25 30  
 Gly Lys Arg Pro Pro Gln Lys Gln Phe Leu Gln Ile Val Pro Arg Gly  
 35 40 45  
 Val Met Phe Asp Gly Gln Thr Gly Leu Ile Lys His Lys Thr Gly Arg  
 50 55 60  
 Leu Pro Leu Met Phe Tyr Arg Glu Ile Lys His Leu Leu Ser His Asp  
 65 70 75 80  
 Met Val Trp Pro Cys Pro Trp Arg Glu Thr Leu Val Gly Arg Val Val  
 85 90 95

Gly Pro Ile Arg Phe His Thr Tyr Asp Gln Thr Asp Ala Val Leu Phe  
 100 105 110  
 Phe Asp Ser Pro Glu Asn Val Ser Pro Arg Tyr Arg Gln His Leu Val  
 115 120 125  
 Pro Ser Gly Asn Val Leu Arg Phe Phe Gly Ala Thr Glu His Gly Tyr  
 130 135 140  
 Ser Ile Cys Val Asn Val Phe Gly Gln Arg Ser Tyr Phe Tyr Cys Glu  
 145 150 155 160  
 Tyr Ser Asp Thr Asp Arg Leu Arg Glu Val Ile Ala Ser Val Gly Glu  
 165 170 175  
 Leu Val Pro Glu Pro Arg Thr Pro Tyr Ala Val Ser Val Thr Pro Ala  
 180 185 190  
 Thr Lys Thr Ser Ile Tyr Gly Tyr Gly Thr Arg Pro Val Pro Asp Leu  
 195 200 205  
 Gln Cys Val Ser Ile Ser Asn Trp Thr Met Ala Arg Lys Ile Gly Glu  
 210 215 220  
 Tyr Leu Leu Glu Gln Gly Phe Pro Val Tyr Glu Val Arg Val Asp Pro  
 225 230 235 240  
 Leu Thr Arg Leu Val Ile Asp Arg Arg Ile Thr Thr Phe Gly Trp Cys  
 245 250 255  
 Ser Val Asn Arg Tyr Asp Trp Arg Gln Gln Gly Arg Ala Ser Thr Cys  
 260 265 270  
 Asp Ile Glu Val Asp Cys Asp Val Ser Asp Leu Val Ala Val Pro Asp  
 275 280 285  
 Asp Ser Ser Trp Pro Arg Tyr Arg Cys Leu Ser Phe Asp Ile Glu Cys  
 290 295 300  
 Met Ser Gly Glu Gly Gly Phe Pro Cys Ala Glu Lys Ser Asp Asp Ile  
 305 310 315 320  
 Val Ile Gln Ile Ser Cys Val Cys Tyr Glu Thr Gly Gly Asn Thr Ala  
 325 330 335  
 Val Asp Gln Gly Ile Pro Asn Gly Asn Asp Gly Arg Gly Cys Thr Ser  
 340 345 350  
 Glu Gly Val Ile Phe Gly His Ser Gly Leu His Leu Phe Thr Ile Gly  
 355 360 365  
 Thr Cys Gly Gln Val Gly Pro Asp Val Asp Val Tyr Glu Phe Pro Ser  
 370 375 380  
 Glu Tyr Glu Leu Leu Leu Gly Phe Met Leu Phe Phe Gln Arg Tyr Ala  
 385 390 395 400  
 Pro Ala Phe Val Thr Gly Tyr Asn Ile Asn Ser Phe Asp Leu Lys Tyr  
 405 410 415  
 Ile Leu Thr Arg Leu Glu Tyr Leu Tyr Lys Val Asp Ser Gln Arg Phe  
 420 425 430

Cys Lys Leu Pro Thr Ala Gln Gly Gly Arg Phe Phe Leu His Ser Pro  
 435 440 445  
 Ala Val Gly Phe Lys Arg Gln Tyr Ala Ala Ala Phe Pro Ser Ala Ser  
 450 455 460  
 His Asn Asn Pro Ala Ser Thr Ala Ala Thr Lys Val Tyr Ile Ala Gly  
 465 470 475 480  
 Ser Val Val Ile Asp Met Tyr Pro Val Cys Met Ala Lys Thr Asn Ser  
 485 490 495  
 Pro Asn Tyr Lys Leu Asn Thr Met Ala Glu Leu Tyr Leu Arg Gln Arg  
 500 505 510  
 Lys Asp Asp Leu Ser Tyr Lys Asp Ile Pro Arg Cys Phe Val Ala Asn  
 515 520 525  
 Ala Glu Gly Arg Ala Gln Val Gly Arg Tyr Cys Leu Gln Asp Ala Val  
 530 535 540  
 Leu Val Arg Asp Leu Phe Asn Thr Ile Asn Phe His Tyr Glu Ala Gly  
 545 550 555 560  
 Ala Ile Ala Arg Leu Ala Lys Ile Pro Leu Arg Arg Val Ile Phe Asp  
 565 570 575  
 Gly Gln Gln Ile Arg Ile Tyr Thr Ser Leu Leu Asp Glu Cys Ala Cys  
 580 585 590  
 Arg Asp Phe Ile Leu Pro Asn His Tyr Ser Lys Gly Thr Thr Val Pro  
 595 600 605  
 Glu Thr Asn Ser Val Ala Val Ser Pro Asn Ala Ala Ile Ile Ser Thr  
 610 615 620  
 Ala Ala Val Pro Gly Asp Ala Gly Ser Val Ala Ala Met Phe Gln Met  
 625 630 635 640  
 Ser Pro Pro Leu Gln Ser Ala Pro Ser Ser Gln Asp Gly Val Ser Pro  
 645 650 655  
 Gly Ser Gly Ser Asn Ser Ser Ser Ser Val Gly Val Phe Ser Val Gly  
 660 665 670  
 Ser Gly Ser Ser Gly Gly Val Gly Val Ser Asn Asp Asn His Gly Ala  
 675 680 685  
 Gly Gly Thr Ala Ala Val Ser Tyr Gln Gly Ala Thr Val Phe Glu Pro  
 690 695 700  
 Glu Val Gly Tyr Tyr Asn Asp Pro Val Ala Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Met Ala His Asn Leu Cys Tyr Ser Thr Leu  
 725 730 735  
 Leu Val Pro Gly Gly Glu Tyr Pro Val Asp Pro Ala Asp Val Tyr Ser  
 740 745 750  
 Val Thr Leu Glu Asn Gly Val Thr His Arg Phe Val Arg Ala Ser Val

755	760	765
Arg Val Ser Val Leu Ser Glu Leu Leu Asn Lys Trp Val Ser Gln Arg 770	775	780
Arg Ala Val Arg Glu Cys Met Arg Glu Cys Gln Asp Pro Val Arg Arg 785	790	795 800
Met Leu Leu Asp Lys Glu Gln Met Ala Leu Lys Val Thr Cys Asn Ala 805	810	815
Phe Tyr Gly Phe Thr Gly Val Val Asn Gly Met Met Pro Cys Leu Pro 820	825	830
Ile Ala Ala Ser Ile Thr Arg Ile Gly Arg Asp Met Leu Glu Arg Thr 835	840	845
Ala Arg Phe Ile Lys Asp Asn Phe Ser Glu Pro Cys Phe Leu His Asn 850	855	860
Phe Phe Asn Gln Glu Asp Tyr Val Val Gly Thr Arg Glu Gly Asp Ser 865	870	875 880
Glu Glu Ser Ser Ala Leu Pro Glu Gly Leu Glu Thr Ser Ser Gly Gly 885	890	895
Ser Asn Glu Arg Arg Val Glu Ala Arg Val Ile Tyr Gly Asp Thr Asp 900	905	910
Ser Val Phe Val Arg Phe Arg Gly Leu Thr Pro Gln Ala Leu Val Ala 915	920	925
Arg Gly Pro Ser Leu Ala His Tyr Val Thr Ala Cys Leu Phe Val Glu 930	935	940
Pro Val Lys Leu Glu Phe Glu Lys Val Phe Val Ser Leu Met Met Ile 945	950	955 960
Cys Lys Lys Arg Tyr Ile Gly Lys Val Glu Gly Ala Ser Gly Leu Ser 965	970	975
Met Lys Gly Val Asp Leu Val Arg Lys Thr Ala Cys Glu Phe Val Lys 980	985	990
Gly Val Thr Arg Asp Val Leu Ser Leu Leu Phe Glu Asp Arg Glu Val 995	1000	1005
Ser Glu Ala Ala Val Arg Leu Ser Arg Leu Ser Leu Asp Glu Val 1010	1015	1020
Lys Lys Tyr Gly Val Pro Arg Gly Phe Trp Arg Ile Leu Arg Arg 1025	1030	1035
Leu Val Gln Ala Arg Asp Asp Leu Tyr Leu His Arg Val Arg Val 1040	1045	1050
Glu Asp Leu Val Leu Ser Ser Val Leu Ser Lys Asp Ile Ser Leu 1055	1060	1065
Tyr Arg Gln Ser Asn Leu Pro His Ile Ala Val Ile Lys Arg Leu 1070	1075	1080

Ala Ala Arg Ser Glu Glu Leu Pro Ser Val Gly Asp Arg Val Phe  
 1085 1090 1095

Tyr Val Leu Thr Ala Pro Gly Val Arg Thr Ala Pro Gln Gly Ser  
 1100 1105 1110

Ser Asp Asn Gly Asp Ser Val Thr Ala Gly Val Val Ser Arg Ser  
 1115 1120 1125

Asp Ala Ile Asp Gly Thr Asp Asp Asp Ala Asp Gly Gly Gly Val  
 1130 1135 1140

Glu Glu Ser Asn Arg Arg Gly Gly Glu Pro Ala Lys Lys Arg Ala  
 1145 1150 1155

Arg Lys Pro Pro Ser Ala Val Cys Asn Tyr Glu Val Ala Glu Asp  
 1160 1165 1170

Pro Ser Tyr Val Arg Glu His Gly Val Pro Ile His Ala Asp Lys  
 1175 1180 1185

Tyr Phe Glu Gln Val Leu Lys Ala Val Thr Asn Val Leu Ser Pro  
 1190 1195 1200

Val Phe Pro Gly Gly Glu Thr Ala Arg Lys Asp Lys Phe Leu His  
 1205 1210 1215

Met Val Leu Pro Arg Arg Leu His Leu Glu Pro Ala Phe Leu Pro  
 1220 1225 1230

Tyr Ser Val Lys Ala His Glu Cys Cys  
 1235 1240

<210> 14  
 <211> 1238  
 <212> PRT  
 <213> herpes simplex

<400> 14

Met Phe Cys Ala Ala Gly Gly Pro Thr Ser Pro Gly Gly Lys Ser Ala  
 1 5 10 15

Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro His Asn Pro Arg Gly Ala  
 20 25 30

Thr Gln Thr Ala Pro Pro Pro Cys Arg Arg Gln Asn Phe Tyr Asn Pro  
 35 40 45

His Leu Ala Gln Thr Gly Thr Gln Pro Lys Ala Pro Gly Pro Ala Gln  
 50 55 60

Arg His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro  
 65 70 75 80

Arg Ser Leu Asp Glu Asp Ala Pro Ala Glu Gln Arg Thr Gly Val His  
 85 90 95

Asp Gly Arg Leu Arg Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu  
 100 105 110

Arg Asp Val Leu Arg Val Gly Pro Glu Gly Phe Trp Pro Arg Arg Leu

115	120	125
Arg Leu Trp Gly Gly Ala Asp His Ala Pro Lys Gly Phe Asp Pro Thr 130 135 140		
Val Thr Val Phe His Val Tyr Asp Ile Leu Glu His Val Glu His Ala 145 150 155 160		
Tyr Ser Met Arg Ala Ala Gln Leu His Glu Arg Phe Met Asp Ala Ile 165 170 175		
Thr Pro Ala Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly 180 185 190		
His Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met 195 200 205		
Asn Lys Ala Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp 210 215 220		
Leu Cys Glu Arg Leu Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser 225 230 235 240		
Phe Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg 245 250 255		
Ala Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Thr Leu Tyr Tyr Arg Val 260 265 270		
Phe Val Arg Ser Gly Arg Ala Leu Ala Tyr Leu Cys Asp Asn Phe Cys 275 280 285		
Pro Ala Ile Arg Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe 290 295 300		
Ile Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys 305 310 315 320		
Pro Gly Arg Gly Asn Ala Pro Ala Gln Pro Arg Pro Pro Thr Ala Phe 325 330 335		
Gly Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala 340 345 350		
Val Glu Gly Ala Met Cys Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe 355 360 365		
Asp Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val 370 375 380		
Ala Glu Arg Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr 385 390 395 400		
Asp Leu Ser Thr Thr Ala Leu Glu His Ile Leu Leu Phe Ser Leu Gly 405 410 415		
Ser Cys Asp Leu Pro Glu Ser His Leu Ser Asp Leu Ala Ser Arg Gly 420 425 430		
Leu Pro Ala Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu 435 440 445		

Leu Ala Phe Met Thr Phe Val Lys Gln Tyr Gly Pro Glu Phe Val Thr  
 450 455 460  
 Gly Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Val Leu Thr Lys Leu  
 465 470 475 480  
 Thr Glu Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly  
 485 490 495  
 Arg Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys  
 500 505 510  
 Arg Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly  
 515 520 525  
 Ile Ile Thr Asp Lys Val Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val  
 530 535 540  
 Ala Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp  
 545 550 555 560  
 Ile Pro Ala Tyr Tyr Ala Ser Gly Pro Ala Gln Arg Gly Val Ile Gly  
 565 570 575  
 Glu Tyr Cys Val Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys  
 580 585 590  
 Phe Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile  
 595 600 605  
 Asn Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr  
 610 615 620  
 Cys Leu Leu Arg Leu Ala Gly Gln Lys Gly Phe Ile Leu Pro Asp Thr  
 625 630 635 640  
 Gln Gly Arg Phe Arg Gly Leu Asp Lys Glu Ala Pro Lys Arg Pro Ala  
 645 650 655  
 Val Pro Arg Gly Glu Gly Glu Arg Pro Gly Asp Gly Asn Gly Asp Glu  
 660 665 670  
 Asp Lys Asp Asp Asp Glu Asp Glu Asp Gly Asp Glu Arg Glu Glu Val  
 675 680 685  
 Ala Arg Glu Thr Gly Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val  
 690 695 700  
 Leu Asp Pro Thr Ser Gly Phe His Val Asp Pro Val Val Val Phe Asp  
 705 710 715 720  
 Phe Ala Ser Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe  
 725 730 735  
 Ser Thr Leu Ser Leu Arg Pro Glu Ala Val Ala His Leu Glu Ala Asp  
 740 745 750  
 Arg Asp Tyr Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val  
 755 760 765  
 Lys Ala His Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp  
 770 775 780

Leu Ala Met Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Thr Pro  
 785 790 795 800  
 Glu Glu Ala Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val  
 805 810 815  
 Cys Asn Ser Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro  
 820 825 830  
 Cys Leu His Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu  
 835 840 845  
 Leu Ala Thr Arg Ala Tyr Val His Ala Arg Trp Ala Glu Phe Asp Gln  
 850 855 860  
 Leu Leu Ala Asp Phe Pro Glu Ala Ala Gly Met Arg Ala Pro Gly Pro  
 865 870 875 880  
 Tyr Ser Met Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu  
 885 890 895  
 Cys Arg Gly Leu Thr Ala Ala Gly Leu Val Ala Met Gly Asp Lys Met  
 900 905 910  
 Ala Ser His Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu  
 915 920 925  
 Cys Glu Lys Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr  
 930 935 940  
 Ile Gly Val Ile Cys Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu  
 945 950 955 960  
 Val Arg Lys Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu  
 965 970 975  
 Val Asp Leu Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala  
 980 985 990  
 Leu Ala Glu Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu  
 995 1000 1005  
 Gly Leu Gln Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg  
 1010 1015 1020  
 Ile Thr Asp Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala  
 1025 1030 1035  
 Glu Leu Ser Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala  
 1040 1045 1050  
 His Leu Thr Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val  
 1055 1060 1065  
 Pro Ser Ile Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr  
 1070 1075 1080  
 Arg Glu Val Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu  
 1085 1090 1095  
 Leu Asp Ala Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala

1100                      1105                      1110  
 Leu Pro Ser Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser His Ala  
 1115                      1120                      1125  
 Asp Pro Pro Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser  
 1130                      1135                      1140  
 Glu Leu Ala Glu Asp Pro Gly Tyr Ala Ile Ala Arg Gly Val Pro  
 1145                      1150                      1155  
 Leu Asn Thr Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys  
 1160                      1165                      1170  
 Val Thr Phe Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu  
 1175                      1180                      1185  
 Ser Leu Leu Lys Arg Phe Ile Pro Glu Thr Trp His Pro Pro Asp  
 1190                      1195                      1200  
 Asp Val Ala Ala Arg Leu Arg Ala Ala Gly Phe Gly Pro Ala Gly  
 1205                      1210                      1215  
 Ala Gly Ala Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala  
 1220                      1225                      1230  
 Phe Asp Thr Leu Ala  
 1235  
 <210> 15  
 <211> 1240  
 <212> PRT  
 <213> herpes simplex  
 <400> 15  
 Met Phe Cys Ala Ala Gly Gly Pro Ala Ser Pro Gly Gly Lys Ser Ala  
 1                      5                      10                      15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro His Asn Pro Arg Gly Ala  
 20                      25                      30  
 Thr Gln Thr Ala Pro Pro Pro Cys Arg Arg Gln Asn Phe Tyr Asn Pro  
 35                      40                      45  
 His Leu Ala Gln Thr Gly Thr Gln Pro Lys Ala Pro Gly Pro Ala Gln  
 50                      55                      60  
 Arg His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro  
 65                      70                      75                      80  
 Arg Ser Leu Asp Glu Asp Ala Pro Ala Glu Gln Arg Thr Gly Val His  
 85                      90                      95  
 Asp Gly Arg Leu Arg Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu  
 100                      105                      110  
 Arg Asp Val Leu Arg Val Gly Pro Glu Gly Phe Trp Pro Arg Arg Leu  
 115                      120                      125  
 Arg Leu Trp Gly Gly Ala Asp His Ala Pro Glu Gly Phe Asp Pro Thr  
 130                      135                      140

Val Thr Val Phe His Val Tyr Asp Ile Leu Glu His Val Glu His Ala  
 145 150 155 160  
 Tyr Ser Met Arg Ala Ala Gln Leu His Glu Arg Phe Met Asp Ala Ile  
 165 170 175  
 Thr Pro Ala Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly  
 180 185 190  
 His Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met  
 195 200 205  
 Asn Lys Ala Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp  
 210 215 220  
 Leu Cys Glu Arg Leu Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser  
 225 230 235 240  
 Phe Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg  
 245 250 255  
 Ala Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Thr Leu Tyr Tyr Arg Val  
 260 265 270  
 Phe Val Arg Ser Gly Arg Ala Leu Ala Tyr Leu Cys Asp Asn Phe Cys  
 275 280 285  
 Pro Ala Ile Arg Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe  
 290 295 300  
 Ile Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys  
 305 310 315 320  
 Pro Gly Arg Gly Asn Ala Pro Ala Gln Pro Arg Pro Pro Thr Ala Phe  
 325 330 335  
 Gly Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala  
 340 345 350  
 Val Glu Gly Ala Met Cys Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe  
 355 360 365  
 Asp Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val  
 370 375 380  
 Ala Glu Arg Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr  
 385 390 395 400  
 Asp Leu Ser Thr Thr Ala Leu Glu His Ile Leu Leu Phe Ser Leu Gly  
 405 410 415  
 Ser Cys Asp Leu Pro Glu Ser His Leu Ser Asp Leu Ala Ser Arg Gly  
 420 425 430  
 Leu Pro Ala Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu  
 435 440 445  
 Leu Ala Phe Met Thr Phe Val Lys Gln Tyr Gly Pro Glu Phe Val Thr  
 450 455 460  
 Gly Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Val Leu Thr Lys Leu

465                      470                      475                      480  
 Thr Glu Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly  
                                  485                      490                      495  
 Arg Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys  
                                  500                      505                      510  
 Arg Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly  
                                  515                      520                      525  
 Ile Ile Thr Asp Lys Val Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val  
                                  530                      535                      540  
 Ala Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp  
 545                      550                      555                      560  
 Ile Pro Ala Tyr Tyr Ala Ser Gly Pro Ala Gln Arg Gly Val Ile Gly  
                                  565                      570                      575  
 Glu Tyr Cys Val Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys  
                                  580                      585                      590  
 Phe Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile  
                                  595                      600                      605  
 Asn Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr  
                                  610                      615                      620  
 Cys Leu Leu Arg Leu Ala Gly Gln Lys Gly Phe Ile Leu Pro Asp Thr  
 625                      630                      635                      640  
 Gln Gly Arg Phe Arg Gly Leu Asp Lys Glu Ala Pro Lys Arg Pro Ala  
                                  645                      650                      655  
 Val Pro Arg Gly Glu Gly Glu Arg Pro Gly Asp Gly Asn Gly Asp Glu  
                                  660                      665                      670  
 Asp Lys Asp Asp Asp Glu Asp Gly Asp Glu Asp Gly Asp Glu Arg Glu  
                                  675                      680                      685  
 Glu Val Ala Arg Glu Thr Gly Gly Arg His Val Gly Tyr Gln Gly Ala  
                                  690                      695                      700  
 Arg Val Leu Asp Pro Thr Ser Gly Phe His Val Asp Pro Val Val Val  
 705                      710                      715                      720  
 Phe Asp Phe Ala Ser Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu  
                                  725                      730                      735  
 Cys Phe Ser Thr Leu Ser Leu Arg Pro Glu Ala Val Ala His Leu Glu  
                                  740                      745                      750  
 Ala Asp Arg Asp Tyr Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe  
                                  755                      760                      765  
 Phe Val Lys Ala His Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg  
                                  770                      775                      780  
 Asp Trp Leu Ala Met Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser  
 785                      790                      795                      800

Pro Pro Glu Glu Ala Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys  
 805 810 815  
 Val Val Cys Asn Ser Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu  
 820 825 830  
 Leu Pro Cys Leu His Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu  
 835 840 845  
 Met Leu Leu Ala Thr Arg Ala Tyr Val His Ala Arg Trp Ala Glu Phe  
 850 855 860  
 Asp Gln Leu Leu Ala Asp Phe Pro Glu Ala Ala Gly Met Arg Ala Pro  
 865 870 875 880  
 Gly Pro Tyr Ser Met Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe  
 885 890 895  
 Val Leu Cys Arg Gly Leu Thr Ala Ala Gly Leu Val Ala Met Gly Asp  
 900 905 910  
 Lys Met Ala Ser His Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys  
 915 920 925  
 Leu Glu Cys Glu Lys Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys  
 930 935 940  
 Lys Tyr Ile Gly Val Ile Cys Gly Gly Lys Met Leu Ile Lys Gly Val  
 945 950 955 960  
 Asp Leu Val Arg Lys Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg  
 965 970 975  
 Ala Leu Val Asp Leu Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala  
 980 985 990  
 Ala Ala Leu Ala Glu Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu  
 995 1000 1005  
 Pro Glu Gly Leu Gln Ala Phe Gly Ala Val Leu Val Asp Ala His  
 1010 1015 1020  
 Arg Arg Ile Thr Asp Pro Glu Arg Asp Ile Gln Asp Phe Val Leu  
 1025 1030 1035  
 Thr Ala Glu Leu Ser Arg His Pro Arg Ala Tyr Thr Asn Lys Arg  
 1040 1045 1050  
 Leu Ala His Leu Thr Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala  
 1055 1060 1065  
 Gln Val Pro Ser Ile Lys Asp Arg Ile Pro Tyr Val Ile Val Ala  
 1070 1075 1080  
 Gln Thr Arg Glu Val Glu Glu Thr Val Ala Arg Leu Ala Ala Leu  
 1085 1090 1095  
 Arg Glu Leu Asp Ala Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro  
 1100 1105 1110  
 Ala Ala Leu Pro Ser Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser  
 1115 1120 1125

His Ala Asp Pro Pro Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu  
 1130 1135 1140  
 Val Ser Glu Leu Ala Glu Asp Pro Gly Tyr Ala Ile Ala Arg Gly  
 1145 1150 1155  
 Val Pro Leu Asn Thr Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala  
 1160 1165 1170  
 Ala Cys Val Thr Phe Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile  
 1175 1180 1185  
 Thr Glu Ser Leu Leu Lys Arg Phe Ile Pro Glu Thr Trp His Pro  
 1190 1195 1200  
 Pro Asp Asp Val Ala Ala Arg Leu Arg Ala Ala Gly Phe Gly Pro  
 1205 1210 1215  
 Ala Gly Ala Gly Ala Thr Ala Glu Glu Thr Arg Arg Met Leu His  
 1220 1225 1230  
 Arg Ala Phe Asp Thr Leu Ala  
 1235 1240  
 <210> 16  
 <211> 1235  
 <212> PRT  
 <213> herpes simplex  
 <400> 16  
 Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala  
 1 5 10 15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala  
 20 25 30  
 Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr  
 35 40 45  
 Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg  
 50 55 60  
 His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg  
 65 70 75 80  
 Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp  
 85 90 95  
 Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg  
 100 105 110  
 Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg  
 115 120 125  
 Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val  
 130 135 140  
 Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr  
 145 150 155 160

Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr  
 165 170 175  
 Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His  
 180 185 190  
 Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn  
 195 200 205  
 Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu  
 210 215 220  
 Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe  
 225 230 235 240  
 Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr  
 245 250 255  
 Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr  
 260 265 270  
 Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro  
 275 280 285  
 Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile  
 290 295 300  
 Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro  
 305 310 315 320  
 Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly  
 325 330 335  
 Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile  
 340 345 350  
 Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp  
 355 360 365  
 Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala  
 370 375 380  
 Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp  
 385 390 395 400  
 Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser  
 405 410 415  
 Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu  
 420 425 430  
 Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu  
 435 440 445  
 Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly  
 450 455 460  
 Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495

Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560  
 Pro Ala Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765  
 Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780  
 Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800  
 Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro Cys Leu His

820	825	830
Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr 835	840	845
Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala 850	855	860
Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met 865	870	875 880
Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly 885	890	895
Leu Thr Ala Ala Gly Leu Thr Ala Met Gly Asp Lys Met Ala Ser His 900	905	910
Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys 915	920	925
Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val 930	935	940
Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys 945	950	955 960
Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu 965	970	975
Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu 980	985	990
Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln 995	1000	1005
Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp 1010	1015	1020
Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser 1025	1030	1035
Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr 1040	1045	1050
Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile 1055	1060	1065
Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val 1070	1075	1080
Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala 1085	1090	1095
Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser 1100	1105	1110
Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser His Ala Asp Pro Pro 1115	1120	1125
Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala 1130	1135	1140

Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr  
 1145 1150 1155  
 Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe  
 1160 1165 1170  
 Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu  
 1175 1180 1185  
 Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Ala  
 1190 1195 1200  
 Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala Gly Ala  
 1205 1210 1215  
 Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr  
 1220 1225 1230  
 Leu Ala  
 1235  
 <210> 17  
 <211> 1235  
 <212> PRT  
 <213> herpes simplex  
 <400> 17  
 Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala  
 1 5 10 15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala  
 20 25 30  
 Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr  
 35 40 45  
 Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg  
 50 55 60  
 His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg  
 65 70 75 80  
 Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp  
 85 90 95  
 Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg  
 100 105 110  
 Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg  
 115 120 125  
 Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val  
 130 135 140  
 Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr  
 145 150 155 160  
 Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr  
 165 170 175  
 Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His

180	185	190
Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn 195 200 205		
Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu 210 215 220		
Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe 225 230 235 240		
Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr 245 250 255		
Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr 260 265 270		
Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro 275 280 285		
Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile 290 295 300		
Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro 305 310 315 320		
Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly 325 330 335		
Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile 340 345 350		
Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp 355 360 365		
Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala 370 375 380		
Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp 385 390 395 400		
Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser 405 410 415		
Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu 420 425 430		
Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu 435 440 445		
Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly 450 455 460		
Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr 465 470 475 480		
Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg 485 490 495		
Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg 500 505 510		

Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560  
 Pro Ala Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Ile Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765  
 Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780  
 Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800  
 Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro Cys Leu His  
 820 825 830  
 Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr  
 835 840 845

Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala  
 850 855 860  
 Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly  
 885 890 895  
 Leu Thr Ala Ala Gly Leu Thr Ala Met Gly Asp Lys Met Ala Ser His  
 900 905 910  
 Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys  
 915 920 925  
 Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val  
 930 935 940  
 Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys  
 945 950 955 960  
 Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu  
 965 970 975  
 Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu  
 980 985 990  
 Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln  
 995 1000 1005  
 Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp  
 1010 1015 1020  
 Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser  
 1025 1030 1035  
 Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr  
 1040 1045 1050  
 Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile  
 1055 1060 1065  
 Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val  
 1070 1075 1080  
 Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala  
 1085 1090 1095  
 Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser  
 1100 1105 1110  
 Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser Pro Ala Asp Pro Pro  
 1115 1120 1125  
 Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala  
 1130 1135 1140  
 Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr  
 1145 1150 1155  
 Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe

1160                      1165                      1170  
 Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu  
 1175                      1180                      1185  
 Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Thr  
 1190                      1195                      1200  
 Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala Gly Ala  
 1205                      1210                      1215  
 Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr  
 1220                      1225                      1230  
 Leu Ala  
 1235  
 <210> 18  
 <211> 1235  
 <212> PRT  
 <213> herpes simplex  
 <400> 18  
 Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala  
 1                      5                      10                      15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala  
 20                      25                      30  
 Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr  
 35                      40                      45  
 Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg  
 50                      55                      60  
 His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg  
 65                      70                      75                      80  
 Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp  
 85                      90                      95  
 Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg  
 100                      105                      110  
 Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg  
 115                      120                      125  
 Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val  
 130                      135                      140  
 Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr  
 145                      150                      155                      160  
 Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr  
 165                      170                      175  
 Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His  
 180                      185                      190  
 Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn  
 195                      200                      205

Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu  
 210 215 220  
 Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe  
 225 230 235 240  
 Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr  
 245 250 255  
 Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr  
 260 265 270  
 Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro  
 275 280 285  
 Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile  
 290 295 300  
 Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro  
 305 310 315 320  
 Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly  
 325 330 335  
 Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile  
 340 345 350  
 Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp  
 355 360 365  
 Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala  
 370 375 380  
 Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp  
 385 390 395 400  
 Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser  
 405 410 415  
 Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu  
 420 425 430  
 Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu  
 435 440 445  
 Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly  
 450 455 460  
 Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495  
 Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala

530	535	540
Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile 545 550 555 560		
Pro Thr Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu 565 570 575		
Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe 580 585 590		
Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn 595 600 605		
Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys 610 615 620		
Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln 625 630 635 640		
Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala 645 650 655		
Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asn 660 665 670		
Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu 675 680 685		
Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro 690 695 700		
Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser 705 710 715 720		
Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu 725 730 735		
Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr 740 745 750		
Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His 755 760 765		
Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met 770 775 780		
Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala 785 790 795 800		
Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser 805 810 815		
Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro Cys Leu His 820 825 830		
Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr 835 840 845		
Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala 850 855 860		

Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly  
 885 890 895  
 Leu Thr Ala Ala Gly Leu Thr Ala Val Gly Asp Lys Met Ala Ser His  
 900 905 910  
 Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys  
 915 920 925  
 Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val  
 930 935 940  
 Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys  
 945 950 955 960  
 Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu  
 965 970 975  
 Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu  
 980 985 990  
 Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln  
 995 1000 1005  
 Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp  
 1010 1015 1020  
 Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser  
 1025 1030 1035  
 Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr  
 1040 1045 1050  
 Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile  
 1055 1060 1065  
 Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val  
 1070 1075 1080  
 Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala  
 1085 1090 1095  
 Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser  
 1100 1105 1110  
 Pro Ala Lys Arg Pro Arg Glu Thr Pro Ser Pro Ala Asp Pro Pro  
 1115 1120 1125  
 Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala  
 1130 1135 1140  
 Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr  
 1145 1150 1155  
 Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe  
 1160 1165 1170  
 Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu  
 1175 1180 1185

Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Ala  
 1190 1195 1200  
 Ala Arg Leu Arg Thr Ala Gly Phe Gly Ala Val Gly Ala Gly Ala  
 1205 1210 1215  
 Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr  
 1220 1225 1230  
 Leu Ala  
 1235  
 <210> 19  
 <211> 1235  
 <212> PRT  
 <213> herpes simplex  
 <400> 19  
 Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala  
 1 5 10 15  
 Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala  
 20 25 30  
 Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr  
 35 40 45  
 Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg  
 50 55 60  
 His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg  
 65 70 75 80  
 Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp  
 85 90 95  
 Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg  
 100 105 110  
 Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg  
 115 120 125  
 Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val  
 130 135 140  
 Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr  
 145 150 155 160  
 Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr  
 165 170 175  
 Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His  
 180 185 190  
 Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn  
 195 200 205  
 Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu  
 210 215 220

Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe  
 225 230 235 240  
 Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr  
 245 250 255  
 Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr  
 260 265 270  
 Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro  
 275 280 285  
 Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile  
 290 295 300  
 Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro  
 305 310 315 320  
 Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly  
 325 330 335  
 Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile  
 340 345 350  
 Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp  
 355 360 365  
 Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala  
 370 375 380  
 Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp  
 385 390 395 400  
 Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser  
 405 410 415  
 Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu  
 420 425 430  
 Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu  
 435 440 445  
 Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly  
 450 455 460  
 Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr  
 465 470 475 480  
 Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg  
 485 490 495  
 Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg  
 500 505 510  
 Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile  
 515 520 525  
 Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala  
 530 535 540  
 Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile  
 545 550 555 560

Pro Ala Tyr Tyr Ala Ala Gly Pro Ala Gln Arg Gly Val Ile Gly Glu  
 565 570 575  
 Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe  
 580 585 590  
 Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn  
 595 600 605  
 Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys  
 610 615 620  
 Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln  
 625 630 635 640  
 Gly Arg Phe Arg Gly Gly Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala  
 645 650 655  
 Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp  
 660 665 670  
 Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu  
 675 680 685  
 Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Arg Val Leu Asp Pro  
 690 695 700  
 Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser  
 705 710 715 720  
 Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu  
 725 730 735  
 Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr  
 740 745 750  
 Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His  
 755 760 765  
 Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met  
 770 775 780  
 Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala  
 785 790 795 800  
 Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser  
 805 810 815  
 Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro Cys Leu His  
 820 825 830  
 Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr  
 835 840 845  
 Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala  
 850 855 860  
 Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met  
 865 870 875 880  
 Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly

885	890	895
Leu Thr Ala Ala Gly Leu Thr Ala Val Gly Asp Lys Met Ala Ser His		
900	905	910
Ile Ser Arg Ala Leu Phe Leu Ser Pro Ile Lys Leu Glu Cys Glu Lys		
915	920	925
Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val		
930	935	940
Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys		
945	950	955
Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu		
965	970	975
Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu		
980	985	990
Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln		
995	1000	1005
Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp		
1010	1015	1020
Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser		
1025	1030	1035
Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr		
1040	1045	1050
Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile		
1055	1060	1065
Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val		
1070	1075	1080
Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala		
1085	1090	1095
Ala Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser		
1100	1105	1110
Pro Ala Lys Arg Pro Arg Glu Thr Pro Leu His Ala Asp Pro Pro		
1115	1120	1125
Gly Gly Ala Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala		
1130	1135	1140
Glu Asp Pro Ala Tyr Ala Ile Ala His Gly Val Ala Leu Asn Thr		
1145	1150	1155
Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala Ala Cys Val Thr Phe		
1160	1165	1170
Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr Glu Ser Leu Leu		
1175	1180	1185
Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp Asp Val Ala		
1190	1195	1200

Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala Gly Ala  
1205 1210 1215

Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp Thr  
1220 1225 1230

Leu Ala  
1235